THE ALBERTA JOURNAL O EDUCATIONAL RESEARC

VOLUME XXXII NUMBER 3 SEPTEMBER 19

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMONT

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

Published in March, June, September, and December by the
Faculty of Education, University of Alberta

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

EDITOR: Walter H. Worth

CONSULTING EDITORS

Reno A. Bosetti,
Alberta Department of Education
Robert C. Bryce,
University of Regina
Robert K. Crocker,
Memorial University of Newfoundland
Kathy G. Dueck,
University of Calgary
Michael G. Fullan,
Ontario Institute for Studies in Education
Myrna L. Greene,
University of Lethbridge
Geoffrey B. Isherwood,
McGill University
Joanne C. Keselman,
University of Manitoba
Terry R. Morrison,
Athabasca University
Alan R. Pence,
University of Victoria
Bikkar S. Randhawa,
University of Saskatchewan
C. Jim Richardson,
University of New Brunswick
W. Todd Rogers,
University of British Columbia
Jaap Tuinman,
Simon Fraser University
Donald J. Willower,
Pennsylvania State University
John D. Wilson,
Moray House College of Education

FACULTY ADVISORY COMMITTEE

Thomas O. Maguire
Erwin Miklos (Chair)
Alton T. Olson
Daiyo Sawada

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

The subscription rate is $24.00 per year for individuals, $32.00 per year for institutions. Add $7.00 for delivery outside of Canada. Single copies are $7.00 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of publication.

Address all communications and manuscript submissions to *The Alberta Journal of Educational Research*, Faculty of Education, Publication Services, 4-116 Education North, University of Alberta, Edmonton, Alberta, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXXII, Number 3 *September, 1986*

CONTENTS

THE EDITOR'S PAGE: *Research Impotence: Blessing or Blight?* 175

LOGO and the Transfer of Problem Solving: An Empirical Test 176
J. Mitterer and L. Rose-Krasnor

Factors Related to Cursive Writing Skill 195
H. Ratzlaff and D. Armitage

Two Teachers' Perspectives on Social Studies Teaching in French Immersion 201
Y.T.M. Mahé

Critical Incidents in the Supervision of Student Teaching: A Cluster Analysis of the Perceptions of Graduate Transfer Students 212
W.B. Boldt and B. Housego

Precocious Readers: A Comparative Study of Early Reading and Non Early Reading Kindergarten Entrants 223
J. Grant and C. Brown

Selected Job Stressors, Emotional Exhaustion, Job Satisfaction, and Thrust Behavior of the High School Principal 234
R.B. Kottkamp and A.L. Travlos

PERSPECTIVES

Making Research Matter 249
E.A. Holdaway

FACULTY OF EDUCATION
The University of Alberta

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September 1986, 175-175*

THE EDITOR'S PAGE

Research Impotence: Blessing or Blight?

Holdaway's article in the Perspectives section of this issue reminds us that much of the research in education has not been a potent source of enlightenment or change. This impotence may be more of a blessing than a blight!

Many research studies are "single-shot" efforts by students seeking a graduate degree, or professors pursuing career advancement. These ventures tend to focus on "targets of opportunity" based on the researchers' idiosyncratic perceptions of significant educational problems, available time, funds, and talent, and access to data. The conceptual frameworks which guide the investigations are not readily apparent. Data collection is usually emphasized more than data analysis. Contributions to theory are rarely assessed. Suggestions for further research, when offered, are seldom followed up either by the authors or the readers.

The result is a "shotgun effect"—a scattering of fragmented information about inconsequential matters. This kind of "buckshot" is such a poor guide to action in the classroom or the boardroom that we ought to feel relief, rather than regret, that it is so seldom used. Otherwise, researchers might be the scapegoats of our educational system and educational research might enjoy even less support than it does now.

Is this too cynical and harsh an indictment? Probably. But it does suggest the desirability of some modification of current research patterns and traditions.

One place to start is with summary or interpretive investigations designed primarily for the scholarly community. These would analyze the rationales of recent research studies about various phenomena and seek to identify ways of overcoming their conceptual weaknesses while building on their strengths. Such analyses could facilitate more meaningful research activities rooted in carefully examined and elaborated conceptual frameworks and planned in relation to the larger educational issues they may potentially influence.

Another is to move beyond superficial triangulation in multimethod research toward interrelating the analysis of quantitative and qualitative data in ways which will increase the trustworthiness of research findings. A 1986 Sage publication, *Linking Data* by Fielding and Fielding, is illustrative of attempts to achieve such meaningful convergence.

Still another starting point is for more groups of professors in education to engage in thematic longer-term collaborative research and to encourage graduate students to relate their theses to these undertakings. This practice is as uncommon in education as it is common in many other disciplines. The outcome could be a critical mass of research evidence which would inform policy and practice to replace the many snippets of unrelated information currently being produced.

Articles reflecting any or all of these approaches would be welcomed, as would one which challenges the views expressed herein.

WHW

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September 1986, 176-194*

JOHN MITTERER

and

LINDA ROSE-KRASNOR
Brock University

LOGO and the Transfer of Problem Solving: An Empirical Test[1]

A study was conducted to investigate the possibility that learning to program in LOGO would result in an improved ability to solve other noncomputer problems. Ninety-six grade six children were divided into four groups (LOGO, BASIC, problem solving, and classroom control). All children were tested on a variety of problem solving tasks before and after training. Experimental groups were given 40 hours of small-group instruction and the children were observed for task-relevant behaviors. These process-oriented observations indicated groups differed in the quality of task-oriented behaviors. Very little LOGO-specific change in problem solving behaviors occurred, however, leading to the conclusion that learning LOGO will not enhance general problem solving skills, at least under the conditions of the current study. Based on these results, it is also concluded that much more theoretical and empirical work will be required before it will be clear just what conditions (if any) will foster the transfer of problem solving skills.

One of the most influential positions concerning how to use computers in schools is that taken by Papert (1980), and his colleagues (Abelson & DiSessa, 1981). Papert offers a vision of the use of computers in schools in which "the child programs the computer and, in doing so, both acquires a sense of mastery over a piece of the most modern and powerful technology and establishes an intimate contact with some of the deepest ideas from science, from mathematics, and from the art of intellectual model building"(p. 5). This vision, as embodied in the computing language LOGO, has captured considerable attention in the educational community. Among the features of Papert's work which have aroused the greatest excitement have been his claims about "powerful ideas." These ideas about thinking are

---

Dr. Mitterer is a cognitive psychologist interested in the applications of cognitive and developmental theories to education.
Dr. Rose-Krasnor is a developmental psychologist with research interests in social interaction during problem solving activities.

encountered while programming LOGO and then become available for use in other intellectual endeavors: "What matters most is that by growing up with a few very powerful theorems, one comes to appreciate how certain ideas can be used as tools to think with over a lifetime" (p. 76).

The educational implications embodied in such claims are far-reaching, mirroring advantages promised in previous years by those in favor of teaching Latin, mathematics, logic, and writing systems for the improvement of general cognitive skills (Pea & Kurland, 1984). For this reason alone, it is important to be confident that such claims rest on solid foundations. Unfortunately, although Papert's ideas possess considerable intuitive appeal, they are not yet fully articulated and, perhaps as a consequence, are not yet well supported either by theoretical arguments or empirical data. This fact, which has led Pea and Kurland (1984) to label Papert's ideas "technoromanticism," has recently led to the emergence of a backlash against the earlier enthusiasm with which LOGO was greeted (e.g., Leousis, 1985), including a call for a moratorium on work with LOGO (Tetenbaum & Mulkeen, 1984). However, it is probably too early to render definitive judgment on such an important position and, as such, it will be more useful to concentrate on articulating the theoretical and empirical issues involved.

Krasnor and Mitterer (1984) and Gallini (1985) have pointed out the relevance of the extensive literature on the transfer of cognitive skills from one domain to another. In essence, Papert (1980) may be interpreted as suggesting that programming is a form of problem solving, that "powerful ideas" are general problem solving heuristics which can be learned by learning to program in LOGO, and that those powerful ideas can be transferred across different types of problems, including those not involving computers. As Krasnor and Mitterer (1984) have shown, the considerable literature concerning problem solving leans to the view that there is no expertise without considerable domain-specific knowledge, and that expertise in one domain does not guarantee expertise in other domains (e.g., Simon, 1980). In this view, it is quite possible that learning to program in LOGO will not foster expertise in other domains. Further, this literature suggests that if general problem solving heuristics are acquired, they tend to be "weak" rather than "strong"; that is, they will generally not be sufficient by themselves for problem solution. Thus even if learning LOGO does result in the appreciation of some powerful ideas, this does not guarantee the manifestation of enhanced general problem solving ability.

Not only do the above arguments reveal little strong theoretical support for Papert's (1980) position, empirical support is still limited as well (see Krasnor & Mitterer, 1984; Pea & Kurland, 1983 for reviews). Krasnor and Mitterer pointed out that most of the evidence available at that time was inconclusive and of an anecdotal nature. Early studies generally suffered from methodological shortcomings such as small sample sizes and/or incomplete designs. While subsequent research on LOGO has generally used improved methods, what little has been directed at the transfer problem has yielded mixed results. Pea and Hawkins (1984) found that learning to program in LOGO did not improve children's ability to plan how to carry out a sequence of chores (see also Pea, 1982). Gorman, Jr. & Bourne (1983) showed that children given an extra hour of individual computer time made significantly fewer errors to criterion on a set of rule-learning problems than did a group given only one half hour extra computer time. Clements and Gullo (1984) found that six-year-olds showed improved reflectiveness, divergent thinking, ability to describe directions, and metacognitive skills as a function of learning LOGO.

In summary, while the position adopted by Papert (1980) is an appealing one, it is not yet well supported by either theoretical arguments or empirical data. The current study was designed to gather evidence relevant to the claim that teaching children to program in LOGO can enhance their ability to solve problems in other domains. It implements several features designed to overcome earlier methodological shortcomings, including a larger sample size, random assignment, a complete design including controls for learning programming and problem solving per se, and a variety of problem solving tasks. Further, this study was deliberately designed to be implemented within the resource limits currently placed on most schools. As such, the results can inform teachers interested in implementing Papert's ideas in today's classrooms.

Method

The main goal of the current study was to investigate the effects of learning LOGO on a variety of social and cognitive tasks. The scope of the investigation was limited to a single age group of 11-12-year-olds and used a "standard" play-oriented version of the LOGO curriculum as described in most published accounts (e.g., Papert, 1980). This age group was selected on the grounds that children who are at the stage of formal operations could best benefit from this instruction (Krasnor & Mitterer, 1984). The effects of LOGO were contrasted with two alternative teaching approaches (one using a computer and one without a computer) and a classroom control condition, using classic problem solving tasks.

One computer-oriented group learned LOGO; the other computer group learned BASIC. A comparison of the BASIC and LOGO groups allows us to detect differences between the effects of learning a structured language (LOGO) and a nonstructured, linear language (BASIC), while controlling for computer exposure. The noncomputer training group received a modification of a well-known problem solving curriculum (the Productive Thinking program). This group was included to help clarify the effects of training which was specifically directed at teaching problem solving without any computer experience. (All children participating in the two noncomputer groups were given instruction in LOGO after the posttesting in order to be fair to all subjects.)

Theoretical analysis of the nature of programming tasks and anecdotal studies led us to predict that the problem solving benefits of learning to program should be apparent in the following problem solving dimensions: (a) efficiency; (b) flexibility in thinking; (c) integration of task-relevant variables; (d) planning and task-orientation; and (e) socioemotional aspects, including attitudinal and motivational components.

Subjects

Four grade six classes in four different elementary schools were involved in the study. Children who were eligible to participate if they and their parents gave permission, were between 11 and 12 years of age and had minimal prior programming experience. From those students meeting the criteria, 24 subjects (12 girls and 12 boys) were randomly selected from each class. Within each class, children were randomly assigned to one of four training conditions: LOGO programming, BASIC programming, noncomputer problem solving, and classroom control. Across the schools, therefore, there were 24 children (12 girls and 12 boys) in each of the four conditions.

The Peabody Picture Vocabulary Test (PPVT) was individually administered prior to the training sessions and provided an index of receptive vocabulary. The four groups did not differ significantly on PPVT scores.

Measures

Pre- and posttest measures were given during the two-week period which preceded and followed the training sessions. All tests were administered by trained testers who were unaware of individual group assignment.

Measures of Problem Solving Efficiency

Efficiency in solving two types of problems was assessed by pre- and posttraining tasks. A block design task was selected to assess the child's ability to analyze visual patterns. Four block design problems (two 4- and two 5-block patterns from the Weschler Intelligence Scale for Children—Revised [WISC-R] test) were individually administered as a pretest measure; four parallel designs were similarly presented as a posttest assessment. The number of seconds to solution (maximum 300 seconds) and number of block placements were recorded as indices of problem solving efficiency. If the design was not completed after five minutes, the tester offered hints until the child was successful.

The child's ability to efficiently solve the Tower of Hanoi problem (Simon, 1975) was also assessed during pre- and posttest periods. This classic task can be recursively solved and taps the child's skill in recognizing repetitive patterns. The four- and five-disk versions were individually presented at pretest and the four- and five-car versions of a Tower isomorph (the parking lot problem) were presented at posttest. The sequence of moves made by the children was recorded by the tester and later arranged in problem-behavior graphs (Newell & Simon, 1972). The child's efficiency of means-end planning was assessed by the number of moves to solution. Two indices of looping structure (tendency to reenter the same problem state again) were derived: (a) the average size of the loops (larger numbers reflect greater deviation from the optimally efficient solution path): and (b) average number of loops which is also inversely related to solution efficiency.

Measures of Problem Solving Flexibility

Problem solving flexibility was measured by two creativity subtests from Wallach and Kogan (1965). Four different designs were presented to each child at pre- and posttesting periods, and children were asked to generate alternate interpretations. Similarly, four common objects were presented at the pretest period and children were asked to generate alternate uses; the procedure was repeated for four different objects at posttesting. Protocols were scored for fluency (number of responses) and novelty (uniqueness) as described by Wallach and Kogan.

Measures of Problem Solving Integration

The child's ability to integrate two variables in problem solving was assessed using two Piagetian proportionality tasks adapted from Siegler (1981). Both tasks were group-administered and scored using Siegler's rule assessment procedure in which the child's answers are classified according to Piagetian stage. Four different stage rules were determined. The first task was a balance scale problem in which

six presentations showed a balance with weights at different distances from the fulcrum. The children were asked to predict which side of the balance would go down if a lever holding the balance in place was removed and to indicate their prediction by circling the chosen alternative on an answer sheet.

The second proportionality task was based on probability judgments. Six problems depicting sets made up of different proportions of white and red marbles were presented. For each problem, the marbles were arranged in two groups, and the children were asked to indicate which group they would want to choose from if they wanted a specified color of marble and had to pick with their eyes closed. Again, answers were recorded on an answer sheet.

Measures of Socioemotional Aspects of Problem Solving, Planning, and Task-Involvement

Measurement of social and emotional components of problem solving included self-report, sociometric, and observational assessments.

Harter's Perceived Competence Scale (1981) was used to assess the child's perceived effectance and mastery motivation in cognitive tasks. This group-administered measure consists of 28 items and produces scores in five effectance dimensions: (a) preference for challenge, (b) curiosity, (c) mastery, (d) judgment, and (e) criteria source. High scores on these scales indicate an internal, mastery-oriented perspective. Children indicated their responses on a four-point scale in a structured alternatives format. The competence scale was given pre- and posttraining.

A computer attitude scale for children (Williams, Coulombe, & Lievrouw, 1983) was used to detect potential changes in attitude during the training period. This scale consists of semantic differential items on seven-point scales representing the dimensions of (a) evaluation, (b) quality, (c) ease of use, and (d) expertness.

The social status of participating children was also monitored to assess changes which may have emerged due to participation in the training conditions. The children's peer reputations were assessed using the Pupil Evaluation Inventory (PEI) developed by Pekarik, Prinz, Liebert, Weintraub, and Neale (1976). On this measure, children were asked to indicate which classmates fit a set of behaviorally oriented descriptions. Factor scores were standardized within each class at pre- and posttesting. The PEI yields factor scores for aggression, likeability, and withdrawal.

The children's social, affective, and task-oriented behaviors were recorded during the tutoring sessions for the three training conditions (LOGO, BASIC, problem solving). Using a time sampling technique, observers focused on each child in randomized order for a 10 second period and then categorized the child's behavior on the following dimensions: (a) task involvement (on, off); (b) affect (positive, negative, neutral); (c) social partner (peer, tutor, peer and tutor, none); (d) information (asks for, receives, gives); (e) attention (asks for, gives to other, receives from other); (f) action (directs others, complies with other's directive, resists other's directive, imitates other's actions, other); (g) overt planning (planning verbalization such as, "Let's do it this way", "We need to get some sticky stuff to hold these together"); and (h) social involvement (solitary or independent activity, equal collaboration, dominant). Category definitions are available from the authors on request.

Observations were made for 20 minutes per session, three days per week, over the seven weeks of training. Three observations per child per session were

collected, totaling a maximum of 54 time samples for each child. Absences and school closures resulted in individual variation in the number of observations. The actual number of time samples obtained averaged 50.8.

Reliability between the two observers was assessed for all 18 children observed in one day per week for 14 weeks. Observers made independent categorizations of each child's behavior. Percentage agreement between the observed ranged from 69.8% to 100% across categories, averaging 91.9%.

Procedure

The training was conducted in two waves, with two schools participating from January to March and the remaining two schools from March to May. In both waves, the pre- and posttest measures were given in the two weeks preceding and following the seven-week training session.

In the three training conditions, children were tutored in groups of six which met each day for a 40-minute period for seven weeks. The children were allowed to choose a partner. Each pair worked together on their own computer or materials, which gave each child 40 minutes of joint experience. A tutor attempted to ensure that each child within a pair had approximately equal computer or materials contact during a session. One day each week, each child was given 20 minutes alone on the computer or materials, while the child's partner worked with the tutor on another training task. Tutors encouraged cooperative work with discovery-oriented processes in all training groups.

The tutors surpervising each group were trained elementary school teachers with programing experience. Each tutor taught all three curricula equally and rotated among the training groups in order to balance teacher-specific effects. The tutors, observers, and supervising researchers met weekly to discuss problems and coordinate the tutor's transition between groups.

As much as possible, the LOGO, BASIC, and problem solving curricula were designed to present similar concepts and provide for equal amounts of peer interaction and individual activity. (Copies of each curriculum can be obtained on request from the researchers.) The LOGO curriculum was adapted from Watt (1983) and Abelson (1981). The BASIC curriculum was drawn heavily from Carlson (1982). The problem solving curriculum was adapted from the Productive Thinking program (Covington, Olton, & Crutchfield, 1974).

Results

All pre- to posttraining changes were assessed using an ANCOVA for posttest scores in which pretest performances were used as covariates.

Problem Solving Efficiency

Efficiency in problem solving was measured by the number of block placements and time to completion on a set of block design tasks. The children's scores were summed across the four pretest problems and the four posttest designs. The average number of block placements was 65.78 for the pretest and 58.32 for the posttest. The ANCOVA for posttest indicated no significant differences among the four experimental groups (group means and standard deviations [SDs] are presented in Table 1). Similarly, the ANCOVA for the number of seconds to completion indicated that experimental groups did not differ on the posttest. Across all groups, $\bar{X} = 465.65$ seconds on the pretest designs and $\bar{X} = 460.99$ seconds on the posttest.

TABLE 1

Modified WIPSI Block Design

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Moves				
Pre	66.7(18.7)	63.2(17.9)	68.0(18.1)	65.3(19.6)
Post	55.4(16.1)	53.5(17.0)	63.0(19.8)	61.4(19.9)
Speed (sec.)				
Pre	460(153)	452(143)	472(150)	479(130)
Post	426(153)	431(198)	510(173)	479(144)

Problem solving effectiveness was also measured on the Tower of Hanoi task by three indices: (a) number of moves; (b) intersection points; and (c) moves per loop. For all indices, scores were summed across the four-and five-item versions. The average number of moves on the pretest Tower version was 122.28 across all four groups; $\bar{X}=109.23$ for the parking lot version given at posttest (group means and SDs are presented in Table 2 for all three indices). Experimental and control groups were not significantly different at posttest, as indicated by the ANCOVA. In addition, groups did not differ on intersection points, although there was an apparent overall difference between pretest performance on the Tower ($\bar{X}=25.61$) and posttest performance on the parking lot problem ($\bar{X}=13.50$).

Moves per loop were also calculated as an inverse index of efficiency. Across all four groups, pretest scores averaged 11.82 moves per loop and posttest scores averaged 11.10. The ANCOVA indicated a significant difference among groups at posttest, $F(3,87)=3.15$, $p<.03$. A comparison of posttest means adjusted for pretest performance indicated that the classroom control group ($\bar{X}=14.49$) and LOGO group ($\bar{X}=12.09$) scores exceeded those shown by the BASIC ($\bar{X}=8.58$) and problem solving ($\bar{X}=9.23$) groups.

Problem Solving Flexibility

Flexibility was measured by fluency (number of responses) and novelty (number of unique responses) for both the patterns and alternate uses tasks. In pre- and posttests, the children's scores were summed across four items. The group means and SDs of the creativity tests are presented in Table 3. Fluency scores on the pattern pretest averaged 12.96 on the pretest and 11.23 on the posttest; for the novelty index, $\bar{X}s=2.90$ and 2.35 for pre- and posttests respectively. The ANCOVAs indicated no significant group differences for novelty, but fluency posttest scores among groups were significantly different, $F(3.87)=3.15$, $p<.013$. The adjusted posttest means indicated that the problem solving training group showed the

TABLE 2

Tower of Hanoi

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Moves				
Pre	131(45.7)	114(39.0)	116(48.0)	128(48.2)
Post	104(36.0)	118(35.9)	106(30.5)	116(45.4)
Intersects				
Pre	28.9(23.5)	21.1(16.5)	23.1(17.0)	29.3(29.8)
Post	11.5(13.6)	13.5(16.5)	12.0(12.8)	17.0(20.0)
Moves/Loop				
Pre	13.7(12.3)	9.18(6.16)	11.5(10.1)	12.9(11.8)
Post	12.4(13.3)	8.13(5.94)	9.18(7.35)	14.7(10.5)

greatest fluency ($\bar{X}=12.74$) followed by the BASIC group ($\bar{X}=11.42$). The LOGO and control groups showed the least fluency ($\bar{X}=10.06$ and 10.70 respectively) on the pattern task.

Across all groups, fluency scores for the alternate uses test averaged 17.10 responses on the alternate uses pretest and 13.11 on the posttest. Novelty scores on this task averaged 1.55 and 1.30 for the pre- and posttest respectively. No significant differences among groups were found on either measure.

Measures of Integration of Task-Relevant Variables

The child's ability to integrate the influence of two variables in a compensatory relationship was assessed with two Piagetian tasks. At formal operations, the child should be able to understand the relationship between distance from the fulcrum and weight in predicting how a balance will respond. Similarly, successful solution of the probability tasks requires the integration of the relative number of two different colored marbles in two arrays. Using Siegler's (1981) presentation and scoring criteria, each child's pattern of responses was categorized as guided by one of four rules. Of the 96 children tested, the responses of 11.5% could not be classified as indicating a consistent rule on the balance problem and 5.2% could not be classified by rule on pre- or posttesting on the probability problem.

In order to assess changes in integration skill due to training, the child's level (rule) on each task on the pretest was subtracted from the level shown on the posttest. The distribution of children who decreased one level, stayed the same, increased one level, and increased two levels is shown in Table 4 for both the balance and probability problems. A chi-square analysis for the balance problem indicated no significant relationship, however, between training group experience and pre- to posttest change. The data shown in Table 4 indicate the lack of variation in

TABLE 3

Creativity Test

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Alternate Uses:				
Fluency				
Pre	17.7(8.8)	17.5(8.0)	16.4(3.4)	16.9(4.8)
Post	13.0(5.8)	13.5(4.4)	12.8(3.9)	13.1(6.9)
Alternate Uses:				
Novelty				
Pre	1.75(1.8)	1.50(3.1)	1.17(1.1)	1.78(2.2)
Post	0.88(.95)	1.17(1.3)	1.46(1.5)	1.67(2.7)
Patterns:				
Fluency				
Pre	12.3(5.5)	12.8(6.6)	11.3(4.2)	15.6(16.8)
Post	9.92(5.5)	11.4(4.1)	12.4(5.6)	11.3(9.2)
Patterns:				
Novelty				
Pre	2.63(2.7)	2.88(3.2)	1.83(1.4)	4.25(8.3)
Post	1.96(2.0)	2.71(2.6)	2.25(2.6)	2.46(4.6)

children's scores on the probability problem. Of those who could be classified by level, all children used rule two at both testing periods. Thus statistical testing would have been meaningless.

Measures of Socioemotional Aspects of Problem Solving

The children's attitudes toward their own effectance in classroom performance were assessed using the Harter Scales (1981). The higher the child's score on these scales, the greater the internality of responses. Four factor scores for each child were calculated at pre- and posttest and are displayed in Table 5. Across all four groups, pre- and posttest scores for the challenge subscale averaged 3.08 and 2.97 respectively, 3.14 and 2.97 for the curiosity subscale, 2.87 and 2.96 for the mastery subscale, 2.28 and 2.49 for the judgment subscale, and 2.71 and 2.95 for the criteria subscale. ANCOVAs for posttest scores indicated groups were not significantly different on any of the five factors.

Potential changes in the children's attitudes towards computers were assessed using a self-report scale with four factor dimensions. Overall pretest scores for the evaluation factor averaged 1.34 on a seven-point scale in which lower scores represent the more positive pole. Posttest scores averaged 2.84. Similarly, pre- and

TABLE 4

Piagetian Tasks

A) Balance Beam

Level (Post-Pre)	Group			
	LOGO	Basic	Problem Solving	Control
−1	2	2	2	2
0	14	14	10	6
1	7	4	8	13
2	0	1	0	0
?	1	3	4	3

B) Probability

Level (Post-Pre)	Group			
	LOGO	Basic	Problem Solving	Control
−1	0	0	0	0
0	23	24	21	23
1	0	0	0	0
2	0	0	0	0
?	1	0	3	1

posttest means for the quality dimensions were 2.03 and 3.18 respectively; 3.44 and 3.64 for ease of use, and 2.37 and 3.33 for expertness of the computer. ANCOVAs for the four posttest factor scores indicated no significant difference among experimental conditions in attitude towards computers. Group means and SDs are presented in Table 6.

The social status of children in participating classrooms was assessed by peer report before and after the training period. Individuals' scores on aggression, withdrawal, and likeability dimensions were standardized within the class, using same-sex peer nominations, at each testing period. ANCOVAs for each of the three dimensions revealed no significant group difference. Table 7 presents group means and SDs.

Observations of the three training groups (LOGO, BASIC, and problem solving) were designed to capture social and affective dimensions of the tutorial experience, as well as overt planning and task involvement. The means and SDs of each of the observations are presented in Table 8.

Because there was a trend toward group differences in number of observations, $F(2,66)=2.61$, $p<.08$, the total number of observations was used as a covariate in the analysis of group differences.

Significant group differences emerged in the social interaction of the children during the training sessions. The number of intervals in which the child was observed interacting with a peer partner, $F(2,65)=7.28$, $p<.001$, teacher partner, $F(2,65=10.84$, $p<.001$, or with both peers and teachers, $F(2,65)=25.90$, $p<.001$,

TABLE 5

Harter Perceived Competence Scale

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Challenge				
Pre	2.91(.66)	2.97(.69)	3.22(.49)	3.20(.59)
Post	2.92(.62)	2.97(.63)	2.90(.54)	3.08(.59)
Curiosity				
Pre	3.18(.47)	3.09(.55)	3.07(.43)	3.21(.51)
Post	2.93(.32)	2.98(.58)	2.82(.43)	3.15(.53)
Master				
Pre	2.85(.59)	2.88(.55)	2.97(.67)	2.77(.68)
Post	2.86(.54)	2.93(.59)	3.03(.48)	2.99(.62)
Judgment				
Pre	2.32(.64)	2.32(.49)	2.31(.63)	2.18(.70)
Post	2.61(.66)	2.42(.64)	2.60(.48)	2.34(.57)
Criteria				
Pre	2.65(.74)	2.82(.75)	2.62(.66)	2.73(.70)
Post	2.95(.53)	2.94(.76)	2.89(.73)	3.03(.86)

showed significant group differences in ANCOVA analysis. Means were adjusted for differences in the total number of observations. Comparisons of adjusted means for peer partners indicated highest peer involvement for the BASIC group ($\bar{X}=15.41$), followed by the LOGO ($\bar{X}=13.39$) and problem solving ($\bar{X}=11.30$) training. Teachers involvement was highest in the LOGO ($\bar{X}=18.65$) group, followed by the BASIC ($\bar{X}=15.64$) and problem solving ($\bar{X}=12.97$) group. The problem solving training, however, showed the greatest number of simultaneous peer and teacher interactions ($\bar{X}=17.95$), compared to the LOGO ($\bar{X}=11.42$) and BASIC ($\bar{X}=10.27$) sessions.

The children in the BASIC group were found to give more information to others ($\bar{X}=15.08$) than children in the other groups ($\bar{X}$s=10.71 for LOGO and 11.83 for problem solving) at a trend level of significance, $F(2,65)=2.83$, $p<.06$.

Groups differed in the directive aspects of interaction. The ANCOVAs for compliance, $F(2,65)=15.66$, $p<.001$) and directs others, $F(2,65)=28.00$, $p<.001$, were significant. Comparison of the adjusted means indicated more compliance in the

TABLE 6

Children's Attitudes Towards Computers

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Evaluation				
Pre	1.33(.60)	1.36(.50)	1.26(.38)	1.41(.43)
Post	2.93(2.4)	2.66(2.1)	2.73(2.2)	3.04(2.3)
Quality				
Pre	2.03(.69)	2.10(.81)	2.05(.87)	1.94(.68)
Post	3.30(1.6)	3.00(1.6)	3.13(1.5)	3.29(1.7)
Ease				
Pre	3.10(1.3)	3.23(1.4)	3.75(1.4)	3.70(1.5)
Post	3.57(1.8)	3.58(1.5)	4.06(1.1)	3.34(1.4)
Expert				
Pre	2.40(1.0)	2.60(1.1)	2.16(.80)	2.31(.75)
Post	3.09(1.7)	3.38(1.7)	3.15(1.6)	3.71(1.8)

TABLE 7

Pupil Evaluation Inventory

Scale	Group			
	LOGO	Basic	Problem Solving	Control
Agression				
Pre	−.34(.86)	−.36(.57)	.10(.94)	.24(1.13)
Post	−.18(.90)	−.43(.45)	.03(1.13)	.12(1.03)
Withdrawal				
Pre	.03(.89)	.14(1.01)	−.32(1.02)	.01(.84)
Post	.11(.79)	.18(1.29)	−.11(.93)	.16(.70)
Likeability				
Pre	.20(.82)	.15(1.03)	−.06(1.13)	−.15(1.01)
Post	.17(.87)	.05(.92)	−.03(1.00)	−.05(.96)

TABLE 8

Observational Measures

Behavior	Group		
	LOGO	Basic	Problem Solving
Off Task	1.25(1.4)	1.21(1.8)	1.88(1.9)
Negative Affect	.75(.85)	.38(.65)	.75(1.29)
Positive Affect	5.92(5.9)	6.17(4.4)	8.96(6.4)
Solitary Activity	8.13(2.6)	8.79(3.1)	8.71(3.1)
Peer Partner	12.8(4.3)	16.0(3.7)	11.3(4.1)
Teacher Partner	18.0(4.5)	16.3(4.9)	12.9(3.8)
Overt Planning	4.00(2.3)	2.75(2.1)	1.83(1.6)
Peer and Adult Partner	10.5(3.5)	11.2(3.0)	17.9(5.2)
Receives Information	23.8(6.0)	23.9(6.2)	25.7(7.2)
Asks For Information	2.67(2.3)	2.96(1.9)	2.96(5.1)
Gives Information	10.7(3.3)	15.1(5.5)	11.8(6.3)
Compliance	2.33(1.3)	3.21(1.4)	0.92(.88)
Directives	2.38(1.3)	4.54(2.8)	0.50(.59)
Resists	0.12(3.4)	0.21(.59)	0.21(.42)
Equal Exchange	4.96(3.9)	4.58(3.5)	4.08(2.2)
Dominant Exchange	36.6(4.9)	39.1(5.3)	38.1(5.7)

LOGO group ($\bar{X}=3.09$) and BASIC ($\bar{X}=2.40$) than in the problem solving group ($\bar{X}=0.98$). Children were also less directive in the problem solving ($\bar{X}=0.57$) than in the BASIC ($\bar{X}=3.58$) and LOGO ($\bar{X}=3.29$) training groups.

The number of time intervals in which direct planning was observed also significantly differed among groups, $F(2,65)=7.59$, $p<.001$. Calculation of adjusted means revealed greater overt planning in the LOGO training group ($\bar{X}=4.58$) than in the BASIC ($\bar{X}=2.13$) and problem solving ($\bar{X}=1.88$) group.

No other significant group differences were found in the ANCOVAs for observational measures.

Discussion

In general, the results of the pre- and posttests indicated relatively few changes in problem solving skills which could be attributed to training. No group differences were found for the ability to integrate two variables. Only one of the efficiency indices showed posttest differences, and that difference favored the problem solving and BASIC programming training. These two groups also showed greater gains in flexibility than did the LOGO and classroom control groups. This rather sparse pattern of results led us to conclude that there was no generalized benefit from 40 hours of computer training on classical measures of problem solving skills.

There were also very few group differences in the attitudinal and mastery measures, indicating that the computer experience did not hold any particular advantage in promoting feelings of internal control over learning among the children in the study. Our subjects were typical grade six children, and our findings certainly do not exclude the possibility that mastery may be enhanced by providing

computer experience to other children, such as those who have failed in traditional school curricula.

Although the focus of this study was the question of general transfer, several process dimensions were measured through observation of the instructional sessions. Our observations focused on social, emotional, and task-oriented components of the children's behavior. We will concentrate on the descriptive rather than explanatory function of this process measurement given that there were few changes which could be attributed to training.

In general, motivation among all three training groups was relatively high. This was reflected in the low levels of off-task behavior and the preponderance of positive over negative affect. The children in the problem solving group apparently had as much fun as the computer group, suggesting that computers in themselves do not have special power to elicit positive affect. In all the training groups there was also a high rate of information exchange (combining categories of giving, receiving, and asking for information). Children in the BASIC group showed more information-giving than the other two groups. The observers' written comments indicated that much of this information was shared with peers. Further observational research should make a finer differentiation in the function of these exchanges (e.g., answering questions, volunteering advice, clarifying ambiguities). Of particular interest would be the extent to which peer tutoring occurs in computer training groups. The potential for major benefits may be relatively strong in this context, possibility leading to increased confidence, sharper communication skills, and enhanced understanding of the concepts being taught.

One of the few positive findings in favor of the LOGO group was a higher frequency of overt planning. Children in this group showed greater explicit verbalization of advance planning than the other groups. These public acts of planning may indicate more reflective consideration of task demands. They may also have served a social function by facilitating shared task-oriented activity. A curriculum which enhances this overt planning between children should also lead to more precise communication skills and cooperative problem solving. These specific processes were not tested in the current study.

One of the initial fears about the widespread use of microcomputers was that children would become "vidiots," isolated from normal peer interaction and deficient in important social experiences. Although the data are sparse (see Rose-Krasnor & Mitterer, 1985 for a review), the research seems to indicate that computers can actually facilitate social interaction when compared to more traditional school activities. The current study clearly revealed a high level of social interaction in both computer training groups. The nature of the interaction was not the same across groups, however. The children in the LOGO group had relatively more interactions with the tutor; the children's interactions in the BASIC group were characterized by relatively more peer-only contact. In contrast, children in the problem solving group showed a relatively high degree of simultaneous contact with peers and tutors. These results were not due to individual differences between tutors, since tutors were counterbalanced across sessions for all groups. Interaction differences may have resulted from differences in the types of problems characteristic of each of the training programs. The desirability of the different patterns of interaction is both an empirical and philosophical question. Our data indicate, however, that there are reliable interaction differences among the groups. Knowledge of these qualitative attributes may help educators decide among training programs once curriculum goals are decided.

Our failure to find significant sex differences was noteworthy. There has been a concern among some educators that girls would show less interest and motivation for computers than boys. Observations indicated that girls were as motivated and involved in the computer training groups as were their male classmates. Our single significant sex difference indicated that boys had a higher level of off-task behavior than girls, but it did not vary by group. Girls and boys showed equivalent affect, planning, information exchange, and social involvement. Similar observations in free choice settings in which other noncomputer activities are available, however, may show a different gender pattern.

The overall pattern of results suggests that the LOGO experience provided in this study did not promote the transfer of LOGO learning to other problem solving situations. This finding, which can be viewed as a disappointing one, raises questions about LOGO transfer and the conditions under which it might be obtained. In searching for those conditions, we confronted a number of important theoretical and methodological issues which will be addressed in the remainder of this paper.

One possibility which must be considered, contrary to Papert's (1980) claims, is that LOGO affords extremely limited opportunities for transfer. Such a view draws support from two major themes of the extensive literature on transfer: that much knowledge is domain-specific and that general problem solving heuristics are very "weak" (e.g., Simon, 1980). The skills which children learn from programming in LOGO may be too domain-specific and/or weak to foster measurable transfer. If this view is correct, then LOGO will be much less generally useful than previously thought. One way to address this possibility is to apply classical principles articulated in the extensive literature on transfer to LOGO learning. Only a few researchers have so far explored this approach (e.g., Krasnor & Mitterer, 1984; Gallini, 1985). Such an analysis might reveal just how much domain-specific knowledge versus general problem solving heuristics are involved in learning LOGO and thus could allow a better understanding of whether or not transfer should be expected.

Another possibility is that LOGO learning will transfer to other domains but not under the conditions of the current study. If this is the case, just what are those other conditions? Our search for alternate conditions has made it clear to us that there exists a large gap between the LOGO philosophy espoused by Papert (1980) and others and "LOGO reality": the actual events which transpire daily in classrooms. Another way of expressing this is that we do not yet have a "theory of LOGO" which is capable of making predictions about the conditions under which transfer might or might not occur. In our view, theoretically motivated positions on the following unresolved key issues (and their interactions) constitute the minimum requirements for such a theory:

(1) What Do Children Learn About LOGO in the First Place?

Proponents of LOGO have so far failed to articulate what it means to learn to program in LOGO. An unstated assumption behind the widespread acceptance of Papert's ideas is that children can and do learn to program in LOGO. The classic pre- posttest design which formed a major part of the current study was adopted with this as an implicit assumption. As such, this design told us little about what children actually learned, how changes occur when learning LOGO, or the specific processes which were responsible for change. A program evaluation using only a before/after design is thus both premature and incomplete; a more process oriented approach is required to yield information about the potential mechanisms of change, as well as the qualitative aspects of children's experiences during training.

Recently, this point has been exceptionally well made by workers at the Bank Street College of Education (e.g. Kurland, 1984; Pea & Kurland, 1984). Pea and Kurland decry the view in which "Programming serves as a 'black box,' an unanalysed activity whose effects are presumed to irradiate those who are exposed to it" (p. 9). Instead they argue that "questions about the development of programming skills require a breakdown of the skills into component abilities, and studies of how specific aspects of programming skills are acquired" (p. 9). While the current study did examine process measures, these were not by and large directly relevant to assessing what was learned.

Thus it will be necessary to focus more attention on defining just what it means to learn to program in LOGO in the first place. We will then have a much better idea of what changes we might reasonably expect to test with pre- posttest designs. For example, Kurland and Pea (1983) recently examined children's ability to learn about recursion—a programming construct which Papert (1980) implied that children can easily learn. However, these investigators convincingly demonstrated that although children can learn to write recursive programs, they do not understand them. On the basis of this result, it may fairly be concluded that testing for the transfer of a recursive style of problem solving is likely to reveal a null result. Such information is obviously useful in determining how to test for transfer. In the course of revealing just what is learned about LOGO, more of such critical analysis and empirical demonstration will make possible better predictions about transfer.

(2) How Much Experience with LOGO is Necessary to Ensure Transfer?

One commonly accepted principle is that transfer will not occur until knowledge is sufficiently well developed in the domain of original learning (Ellis, 1965). Thus how much experience with LOGO will be required before transfer can be expected? In the current study, one obvious possibility is that the children were not given enough learning time to manifest transfer, even though 40 hours of instruction is more than most other training studies have given. However, the LOGO literature does not yet provide any basis for making reasoned estimates of how much experience is enough. Until such time as theory has been developed to the point of allowing principled predictions, we must rely on the results of experiments such as the present one to empirically address the question.

(3) What Type of LOGO Curriculum is Required to Support Learning for Transfer?

It will also be necessary to specify just what type of learning experience is best suited to promote LOGO learning and transfer. Thus it is also possible that our curriculum was inappropriate, even though we deliberately modeled it to reflect the "LOGO philosophy." Perhaps we created too much structure to allow the free exploration necessary to foster the emergence of deep insights into mathematics, computation, and thinking itself. Alternately, perhaps we offered too little structure, thus not ensuring adequate exposure to and practice with critical concepts. Further, Krasnor and Mitterer (1984) noted that the approach to teaching LOGO presented by Papert (1980) failed to consider several well established transfer principles.

An analysis of the existing transfer literature might suggest ways of modifying the LOGO curriculum to make students explicitly aware of the general utility of

powerful ideas. This awareness may enhance the likelihood that learning to program will produce benefits in general problem solving skills. Again the extant literature on LOGO offers only the broad outlines of a LOGO curriculum which might foster transfer. We need much more theoretical and empirical elaboration of this question.

(4) Can Every Child Learn LOGO for Transfer?

Common sense tells us that not every child can learn LOGO well enough to show transfer. Some children, especially the very young, might not be able to learn it at all. Thus it might be argued that our choice of children was not optimal, although it was motivated by the argument that only formal operational children could benefit enough from learning LOGO to show transfer. Perhaps this argument is wrong. Perhaps some other age group (or distribution of males and females, or socioeconomic group, etc.) would have shown transfer. Again, the existing literature on LOGO is simply not adequate enough to support more precise predictions.

(5) What is the Right Way to Measure LOGO Transfer?

It might be argued that the failure of the current study to find general benefits of LOGO stem from the inadequacy of the measures of transfer, even though they are traditionally accepted measures of the efficiency, flexibility, and integration of general problem solving skills. It could be asserted, for example, that these measures do not assess the type of learning that is uniquely occurring when children learn LOGO. A thorough analysis of what is learned from LOGO would be an extremely useful guide for selecting more sensitive and "LOGO-appropriate" measures of transfer.

In retrospect, it is amazing how little we know about the impact of computers and LOGO on children's learning. It will, in our view, take quite some time before we develop a sufficiently rich theoretical base to offer definitive answers to the key questions outlined above. The current study is important in this broad context because it sets boundary conditions on when transfer might be expected by revealing at least one set of conditions under which transfer was *not* obtained. Such studies, each exploring different combinations of experimental parameters and measures, will serve to inform future theoretical work and thus, ultimately, to clarify the usefulness of LOGO for developing general thinking skills.

While this study obviously does not constitute definitive evidence against the existence of some as yet unknown set of conditions under which transfer might occur, it does carry strong practical implications. The seven weeks of daily small-group training by specially trained tutors exceeds the intensity of computer instruction currently available in most schools. Therefore, it is likely that educators realistically should not expect dramatic changes in thinking skills as a result of current levels of computer instruction.

In conclusion, we did not find the revolution in thinking skills which might have been expected on the basis of glowing anecdotal reports and early theoretical claims. As such, we are sceptical that teaching LOGO (or BASIC), especially under current conditions, will be useful *as a way of teaching general thinking skills*. We do, however, acknowledge the possibility that once we know much more about what it means to learn to program and about how to build LOGO curricula to explicitly foster transfer, LOGO might still be an excellent means to this admirable instructional end. Such a demonstration, however, is not likely to become available in the

immediate future. This pessimistic assessment must not, however, be construed as an argument against the use of LOGO in schools. We believe that the LOGO philosophy currently represents one of the most interesting ways of thinking about how to use computers in schools. There are a variety of reasons why LOGO should still be seen as having considerable promise in the field of education. It is, in our opinion, the best language currently available for introducing children to programming *as a skill in its own right*. It is an intrinsically interesting activity, and is also a useful tool for the explicit exploration of topics in mathematics (Abelson & DiSessa, 1981), linguistics (Rowe, 1976), and artificial intelligence (Bundy, 1980).

Notes

1. We would like to express our appreciation to the teachers and principals of the Lincoln County Roman Catholic Separate School Board for their tremendous help and support. The skills of Arlene High and Kathy Bogdan, as research assistants, were critical to the project, and Joan Woodward and Cheryl Wylie's careful observations were very much appreciated. We would also like to especially thank Mary Lou Orr and Roberta Dyment who worked very hard to provide the children with excellent instruction, and to the parents and children who showed great goodwill and enthusiasm. This project was funded by Grant #410-83-0809 from the Social Sciences and Humanities Research Council of Canada. Reprints are available from Linda Rose-Krasnor, Department of Psychology, Brock University, St. Catharines, Ontario, L2S 3A1.

References

Abelson, H. (1981). *LOGO*. New York: McGraw-Hill.

Abelson, H., & DiSessa, A. (1981). *Turtle geometry: The computer as a medium for exploring mathematics*. Cambridge, MA: MIT Press.

Bundy, A. (Ed.). (1980). *Artificial intelligence: An introductory course*. Edinburgh: Edinburgh University Press.

Carlson, E.H. (1982). *Kids and the Apple*. Reston, VA: Reston.

Clements, D.H., & Gullo, D.F. (1984). Effects of computer programming on young children's cognition. *Journal of Educational Psychology, 76*, 1051-1058.

Covington, M., Olton, R., & Crutchfield, R. (1974). *The productive thinking program* (2nd ed.). Columbus, OH: Merrill.

Ellis, H. (1965). *The transfer of learning*. New York: Macmillan.

Gallini, J.K. (1985, February). Instructional conditions for computer-based problem solving environments. *Educational Technology*, 7-11.

Gorman, H. Jr., & Bourne, L. (1983). Learning to think by learning LOGO: Rule learning in third-grade programmers. *Bulletin of the Psychonomic Society, 21*, 165-167.

Harter S. (1981). A new self-report scale of intrinsic versus extrinsic orientation in the classroom: Motivational and informational components. *Developmental Psychology, 17*, 300-312.

Krasnor, L., & Mitterer, J. (1984). LOGO and the development of general problem solving skills. *The Alberta Journal of Educational Research, 30*(2), 133-144.

Kurland, M. (Ed.). (1984). *On the cognitive effects of learning computer programming: A critical look* (Tech. Rep. No. 9). New York: Bank Street College of Education, Center for Children and Technology.

Kurland, M., & Pea, R. (1983). *Children's mental models of recursive LOGO programs* (Tech. Rep. No.10). New York: Bank Street College of Education. Center for Children and Technology.

Leousis, E. (1985, February). Black sheep and LOGO. *Computers in Education*, 7-9.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas.* New York: Basic Books.

Pea, R. (1982). What is planning development the development of? In D. Forbes & M. Greenberg (Eds.), *New directions in child development: The development of planful behavior in children.* San Francisco: Jossey-Bass.

Pea, R., & Hawkins, J. (1984). *Children's planning processes in a chore-scheduling task* (Tech. Rep. No. 11). New York: Bank Street College of Education, Center for Children and Technology.

Pea, R., & Kurland, M. (1983). *LOGO programming and the development of planning skills* (Tech. Rep. No. 16). New York: Bank Street College of Education, Center for Children and Technology.

Pea, R., & Kurland, M. (1984). *On the cognitive effects of learning computer programming: A critical look* (Tech. Rep. No. 9). New York: Bank Street College of Education, Center for Children and Technology.

Pekarik, E., Prinz, R., Liebert, D., Weintraub, S., & Neale, J. (1976). The Pupil Evaluation Inventory: A sociometric technique for assessing children's social behavior. *Journal of Abnormal Child Psychology, 4,* 83-97.

Rose-Krasnor, L., & Mitterer, J.O. (1985). Computer possibilities for preschool education. *The Canadian Journal of Research in Early Childhood Education, 1,* 90-94.

Rowe, N. (1976). Grammar as a programming language. *LOGO MEMO #39.* Cambridge, MA: MIT Artificial Intelligence Laboratory.

Siegler, R. (1981). Developmental sequences within and between concepts. *Monograph of the Society for Research in Child Development, 46,* (Serial No. 189).

Simon, H. (1975). The functional equivalence of problem solving skills. *Cognitive Psychology, 7,* 268-288.

Simon, H. (1980). Problem solving in education. In D.T. Tuma & F. Reif (Eds.), *Problem solving and education: Issues in teaching and research.* Hillsdale, NJ: Erlbaum.

Tetenbaum, T., & Mulkeen, T. (1984, November). LOGO and the teaching of problem solving: A call for a moratorium. *Educational Technology,* 16-19.

Wallach, M., & Kogan, N. (1965). *Modes of thinking in young children.* New York: Holt, Rinehart & Winston.

Watt, D. (1983). *Learning with LOGO.* New York: McGraw-Hill.

Williams, F., Coulombe, J., & Lievrouw, L. (1983). Children's attitudes toward small computers: A preliminary study. *ECTJ, 31,* 3-7.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 195-200.*

HAROLD RATZLAFF
University of British Columbia

and

DOREEN ARMITAGE
Vancouver School District

Factors Related to Cursive Writing Skill

This study explored the relationship between quality of writing and certain instructional strategies and materials. A representative sample (n=137) of grade three students was used to obtain both printing and writing samples during the same school year. A previously developed Diagnostic Inventory was used to evaluate each sample for neatness and the commonly recognized major elements of form. Several relationships were discovered between quality of printing and certain instructional practices. These relationships tended to disappear as the students moved from printing to writing. The results were thought to have implications for the preservice and inservice training of primary grade teachers.

Manuscript writing (i.e., printing) normally gives way to cursive writing in grade two or grade three (Freeman, 1958; Herrick & Okada, 1963; Peck, Askov, & Fairchild, 1980; Renaud & Graff, 1966; Suen, 1975). Niemann's (1971) research suggested that children in the second or third grade make a smoother transition than do older children. Handwriting manuals usually recommend flexibility in the choice of transition time frequently based on an acceptable level of proficiency in manuscript writing (Suen, 1975). Teachers normally attempt to individualize the transition by initially teaching cursive writing to only the better printers, assuming that proficiency in printing is related to proficiency in handwriting. Some research (Otto & Smith, 1980; Wallace & Kauffmann, 1978) supports this latter notion by suggesting that cursive writing be withheld from students who experience difficulty with printing.

Armitage and Ratzlaff (1985), using the *Diagnostic Inventory* as a measure of writing achievement, demonstrated a low or no correlation on six of seven aspects of writing skill when comparing manuscript writing samples and cursive writing

---

Dr. Ratzlaff is an Assistant Professor in the Department of Educational Psychology and Special Education with interest in the measurement and evaluation of teaching and learning.

Ms. Armitage is head teacher in a diagnostic centre with interests in the diagnosis and remediation of learning disabilities.

samples of grade three students. The only aspect that registered a high degree of relationship was neatness. This result caused the researchers to conclude that teachers need not individualize the transition from manuscript to cursive writing and to recommend that all class members be moved from one to the other simultaneously.

The present study sought to explore antecedent factors that might be related to quality of writing, either manuscript or cursive. What practices and/or materials are related to writing achievement as measured by the *Diagnostic Inventory*? Briggs (1970, 1980) demonstrated that students who write badly suffer lower marks when content in written composition is held constant. The study suggests that evaluators tend to form an image of student competency and personality based on quality of writing. There exists, then, a need to promote quality writing and a need to determine if certain teacher practices or materials are related to quality of writing.

Rubin and Henderson (1982) collected normative data on the instructional practices of teachers who taught nine- and ten-year-olds from a borough in London, England. Teachers appeared to have no consistent guidelines for teaching handwriting. In fact, one in ten teachers did not even provide specific handwriting instruction. Among those that did, Rubin and Henderson found much variation in terms of such factors as method of instruction, frequency of practice, length of lesson, and writing style taught. They go on to relate success in handwriting to the motivation and ability of the child, as well as the methods and ability of the teacher.

Method

The sample (n=175) for this study represented seven intact grade three classes from six schools in three school districts. All classes, schools, and districts were chosen so as to obtain a representative sample of grade three students. Data analysis involved those subjects (n=137) from whom both manuscript and cursive writing samples were received.

These students produced the manuscript writing sample in October and the cursive writing sample the following April. The samples were forwarded to the researchers for scoring using the *Diagnostic Inventory* to score seven facets of writing evaluation: letter formation, letter size, uniformity of slant, spacing of letters, spacing of words, word alignment, and neatness. Five raters were used in scoring the samples. These scores, along with responses from a questionnaire completed by the teachers in April, were used in the data analysis.

Results

The reliability and validity of the *Diagnostic Inventory* were established and are reported elsewhere (Armitage & Ratzlaff, 1985). Inter-rater reliability was good and a subject's score was taken as the arithmetic mean of the five raters. This average score was accepted as a reasonable approximation of each subject's true score (Allen & Yen, 1979).

Responses from the teacher questionnaire were studied to determine which instructional strategies and materials were commonly used. Table 1 reveals that approximately 80% of the students received writing instruction that (a) did not involve alphabetic order of presentation, (b) did go from lower case to upper, and

(c) did not mix lower and upper case in a given lesson. No students were instructed to write using vertical strokes before moving to slant.

Table 1 also suggests that teachers were relatively split in the use of special methods. About as many students as not received instruction using (a) hand and arm air movements, (b) verbalization of hand and arm movements, and (c) practice on a chalkboard. Most students (70%) did not trace letters and words on paper prior to writing freehand.

TABLE 1

Comparative Use of Various Instructional Strategies for the Teaching of Writing (n=137)

Instructional Practice	Yes	No	Missing
1. Sequence of Instruction			
a. alphabetic order	25	94	18
b. lower to upper	94	25	18
c. mixed lower and upper	25	94	18
d. vertical to slant	0	119	18
2. Special Methods			
a. air movements (hand)	66	71	0
b. verbalizing movements	73	64	0
c. practice on chalkboard	76	61	0
d. tracing on paper	41	96	0

The main analysis involved the relationship between printing or writing scores and various instructional practices and materials. Table 2 identifies those instructional practices for which there were significant relationships and those for which there were none. Generally, there were no statistically significant (α=.05) relationships between instructional practices or materials and assigned scores for cursive writing samples. Specifically, only scores for letter size (W2), alignment (W6), and neatness (W7) were related to instructional practice for three of nine practices. Higher scores were registered by those who (a) used air movements, (b) traced on paper, and (c) used paper that was not interlined or was 3/8 interlined. A higher score was also registered in neatness (W7) for those students involved in verbalizing the hand movements.

Printing scores, on the other hand, were significantly related on all seven measures for five of nine practices. The other four of nine practices were related on one to four of the measures. Specifically, higher scores were registered by those students whose teachers (a) used alphabetic order in lesson presentations, (b) did not go from lower to upper case but, rather, mixed the order, (c) used hand and arm movements, but (d) did not verbalize these movements. Those students exposed to chalkboard practice tended to score higher on letter size (P2), uniformity of slant (P3), spacing of letters (P4), and alignment (P6). Students using noninterlined paper also scored higher on P2 and P6, as well as letter formation (P1). If students

TABLE 2

Significant* Relationships (X) Between Select Instructional Practices and Facet Scores Assigned to Printing (P) and Writing (W) Samples (N=137)

Instructional Practice	Evaluation Facet													
	P1	P2	P3	P4	P5	P6	P7	W1	W2	W3	W4	W5	W6	W7
1. Instructional Sequence:														
a. alphabetic order	X	X	X	X	X	X	X							
b. lower to upper	X	X	X	X	X	X	X							
c. mixed lower to upper	X	X	X	X	X	X	X							
2. Special Methods:														
a. air movements	X	X	X	X	X	X	X		X				X	X
b. verbalizing movements	X	X	X	X	X	X	X							X
c. chalkboard practice	X	X	X			X								
d. tracing on paper	X				X		X		X				X	
e. type of paper	X	X				X			X				X	X
f. followed pub. series					X									

* $p<.05$

were exposed to tracing on paper, they scored significantly higher on letter formation (P1), spacing of words (P5), and neatness (P7). Students of teachers who used a published series of writing lessons scored higher on one facet only, namely, spacing of words (P5).

Discussion

Certain instructional strategies and materials appear to be related to quality of writing produced by students, especially for printing, thus supporting the research results reported by Rubin and Henderson (1982). Instruction that introduced printing alphabetically and did not routinely move from lower to upper case, rather mixed the order, produced significantly higher scores on all seven evaluated facets. Higher quality printing appeared to be related to instruction involving hand and arm movements alone without a verbalization of the movements. The significant differences or relationships disappeared when cursive writing scores were considered, except for the use of air movements where three of the seven evaluative facets registered significantly higher scores.

Instruction involving practice on the chalkboard or tracing letters and words on paper appeared to be related to higher quality printing, but on mutually exclusive evaluative facets. Taken together, they were related to significantly higher scores on all seven facets. Type of paper used for instruction did not appear to be too significant and resulted in higher scores for both printing and writing on only three of seven evaluative facets. However, as with the use of hand and arm air movements, both tracing on paper and type of paper used appeared to relate to higher scores for writing, not printing alone.

It would be worth noting that it would be rather naive to suggest that the above relationships were unaffected by the instruction received by these grade three students in grade one and two and were solely due to instruction received from their

grade three teacher. Data was not collected on instructional practices and materials used in their grade one and grade two experience.

Conclusion

Several factors investigated in this study appear to be related to quality of printing and writing, particularly the former. However, the relationship between several instructional strategies and materials and quality of printing appears to disappear when the same subjects switch to cursive writing. Sequence of instruction factors had no continuing effect, while several special methods of instruction seemed to have a minor relationship to both manuscript and cursive writing. Based, then, on this study, it would appear that there is no abiding relationship between certain instructional techniques or materials and quality of handwriting.

This research result, along with that of Rubin and Henderson (1982), has implications for the preservice and inservice training of primary grade school teachers. Certain teaching practices, strategies, and materials appear to be related to learning manuscript writing. This relationship would appear to be "washed out" when moving to cursive writing. Based on this research result and that reported earlier (Armitage & Ratzlaff, 1985), it would seem reasonable to suggest that primary teachers move their students simultaneously from printing to writing and not concern themselves too heavily with the issue of identifying the best practices, strategies, and materials.

References

Allen, M.J., & Yen, M. (1979). *Introduction to measurement theory*. Monterey: Brooks/Cole.

Armitage, D., & Ratzlaff, H.C. (1985). The non-correlation of printing and writing skills. *Journal of Educational Research, 78*(3), 174-177.

Briggs, D. (1970). The influence of handwriting on assessment. *Educational Research, 13*, 50-55.

Briggs, D. (1980). A study of the influence of handwriting using examination scripts. *Educational Review, 32*(2).

Freeman, F.N. (1958). The transition from manuscript to cursive writing. *Elementary English, 35*, 366-372.

Herrick, V.E., & Okada, N. (1963). The present scene practices in the teaching of handwriting in the US—1960. In V.E. Herrick (Ed.), *New horizons for research in handwriting*. Madison, WI: University of Wisconsin Press.

Niemann, A.S. (1971). Handwriting, spelling, and creative writing in the primary grades. *Elementary English, 48*, 963-969.

Otto, W., & Smith, R.J. (1980). *Corrective and remedial teaching* (3rd ed.). Boston: Houghton Mifflin.

Peck, M., Askov, E.N., & Fairchild, S.H. (1980). Another decade of research in handwriting: Progress and prospect in the 1970's. *The Journal of Educational Research, 73*, 283-298.

Renaud, A.J., & Graff, P.J. (1966). Parents' opinions about handwriting styles. *Elementary English, 43*, 873-876.

Rubin, N., & Henderson, S.E. (1982). Two sides of the same coin: Variations in teaching methods and failure to learn to write. *Special Education: Forward Trends (Research Supplement), 9*(4), 17-24.

Suen, C.Y. (1975). Handwriting education in Canada. *Saskatchewan Journal of Educational Research and Development, 5*, 46-52.

Wallace, G., & Kauffman, J.M. (1978). *Teaching children with learning problems* (2nd ed.). Columbus, OH: Charles E. Merrill.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 201-211.*

YVETTE T.M. MAHÉ
University of Alberta

Two Teachers' Perspectives on Social Studies Teaching in French Immersion

Researchers have paid little attention to the question of how French immersion teachers manage to integrate second language teaching with social studies teaching. A case study of the perspectives of a grade one and a grade two teacher on social studies teaching in a French immersion situation disclosed that the teachers traded off social studies concept and inquiry skill development goals for second language development goals such as developing students' vocabulary and grammatical structures. The teachers' decisions appeared to be influenced by their perceptions of parents', administrators', and consultants' expectations that they develop students' proficiency in French combined with their own beliefs and assumptions about second language teaching and learning and child socialization.

Since the inception of French immersion programs in the 1960s, experimental studies focusing on the educational outcomes have consistently demonstrated that French immersion schooling has no negative effects on students' academic and linguistic competencies, and on their intellectual and general development (Barik & Swain, 1976; Calgary Catholic Schools, 1979; Council of Ministers of Education, Canada, 1982; Cummins, 1976, 1983; Day & Shapson, 1985; Genesee, 1984; Gray & Cameron, 1975; Laberge, 1984; Lapkin & Swain, 1984; Shapson & Kaufman, 1976, 1978; Stern, 1984; Swain & Lapkin, 1981, 1983). Some researchers attribute the success of immersion to a "communicative competence" methodology, which is supposed to provide students "with a great deal of comprehensible input" (Krashen, 1984, p. 41) so that they learn French incidentally while dealing with the regular school curriculum (Wesche, 1984). However, Bibeau (1984) points out that we still have no significant data on teaching methods or materials actually used by teachers in immersion (p. 46). With respect to social studies teaching, research has not shed light on the kind of social studies knowledge teachers select and organize for early elementary French immersion students.

This article reports the major findings of a case study of two early elementary teachers' perspectives on social studies teaching in French immersion (Mahé, 1984), and discusses the implications of the findings in the light of recent trends in

---

Dr. Mahé is an Assistant Professor of Education in Faculté Saint-Jean responsible for French immersion and social studies teacher preparation.

French immersion research. The argument advanced in this article is that unless teachers are encouraged to develop a commitment to the purpose and goals of social studies within the immersion context, and unless a realistic social studies program is developed for early elementary French immersion which goes beyond the translation of English language documents, immersion students' opportunities to develop social studies concepts and higher order thinking skills, such as inferential reasoning and problem solving abilities, may be limited.

A current trend in research on teaching is based on the assumption that in order to understand teaching, teachers' goals, judgments, and decisions must be understood in relation to their behavior in the classroom context (Shavelson & Stern, 1981, p. 455). Qualitative or naturalistic studies, including case studies, portrayals, participant observation, and ethnographic studies have yielded rich insights, ideas, and hypotheses suggesting that teacher decision making is influenced by a wide variety of external and internal variables (Delamont, 1976; Esland, 1971; Nelson & Cornbleth, 1982; Shavelson & Stern, 1981; Shaver, Davis, & Helburn, 1979; Smith & Geoffrey, 1968; Wilson, 1977; Woods, 1985).

Theoretical and Methodological Framework

A symbolic interactionist perspective guided the researcher in her attempt to describe and explain how teachers' perceptions of the French immersion situation influences their social studies knowledge selection and organization. Symbolic interactionists assume that the way individuals act toward things and others in situations depends on how they define and interpret a variety of external and internal elements constituting their situation (Blumer, 1969; Hewitt, 1979; Mead, 1934).

The case study method. The theoretical dictates of symbolic interactionism require that an investigation has to take place in a natural setting, and that the researcher has to come to grips with the participants' definitions and interpretations of their situation (Blumer, 1969; Hewitt, 1979). The case study method was selected for probing into the constructs underlying a grade one and a grade two teachers' social studies decisions and actions, as it allows a researcher to carry out an intensive in-depth naturalistic investigation of persons in a specific situation (Black & Champion, 1976, pp. 90-91; Helmstadter, 1970, p. 51), and to get close to the participants in order to record their experiences from their point of view (Denzin, 1970, p. 452).

Naturalistic investigators suggest gathering data from different vantage points by using two or more complementary methods, and then triangulating or pitting against each other the data gathered from different methods in order to cross-check and verify the credibility and dependability of the data findings (Denzin, 1970, pp. 459-460; Guba, 1978; Pohland, 1972, p. 11). A limited classroom participant observation methodology, triangulated with unstructured and structured interviewing and document analysis was selected for gathering data and for generating and testing hypotheses, ideas, and insights representative of the two volunteer teachers' perspectives.

Data Gathering Techniques and Procedures

The case study took place in the greater Edmonton, Alberta region. During the 1980-1981 school term, intermittent field work was carried out for four months at the grade one level, and for six months at the grade two level. The grade one

teacher terminated the research project because of a pregnancy leave. The grade two teacher discontinued teaching social studies in April in order to prepare her students for the school board standardized tests.

It was possible for the researcher to observe, at the grade one level, five social studies lessons for a total of 420 minutes, and to observe, at the grade two level, 17 social studies lessons for a total of 843 minutes. During the classroom observations, the writer attempted to minimize her influence on the on going classroom activities by sitting in a corner of the room and recording in shorthand the verbal and nonverbal classroom interactions, and noting how the teachers and students interacted with different types of resource materials.

Interviewing, according to Pohland (1972), complements observing and provides an evidential base which closely approximates the respondent's reality (p. 12). The researcher interviewed the grade one teacher in the privacy of her classroom, generally at the end of the school day, on 12 occasions for a total of 370 minutes, and interviewed the grade two teacher in an unoccupied resource room during the lunch hour on 10 occasions for a total of 287 minutes. With the teachers' permission, the interviews were recorded on a cassette and shorthand field notes were kept on the key interview topics, the key responses, and of the documents examined. Following the interviews, the tapes and field notes were transcribed by the writer in typewritten form. The interview transcripts were given to the teachers for verification as to their accuracy and for generating additional data. The grade one teacher verified 84 pages of interview transcripts, and the grade two teacher verified 109 pages of interview transcripts.

At the beginning of the project, the interview questions were generally unstructured in order to allow the teachers to tell their own stories in their own words (Blumer, 1969; Hewitt, 1979, p. 260; Li, 1981, pp. 45-46). Clarification questions, based on the teachers' previous statements, or questions based on the investigator's classroom observations were asked when it appeared that the teachers had exhausted a topic (Li, 1981, pp. 45-51). As the project evolved, the writer structured her questions in order to test or verify her data interpretations and hypotheses with the teachers.

Data Analysis Techniques and Procedures

The perspective of symbolic interactionism assisted the researcher in identifying, from the data generated from different vantage points, categories of internal and external elements which appeared to influence the two teachers' social studies decisions and actions. The data from the interview and classroom observation transcripts, and the document examination field notes were analyzed, categorized, and synthesized by the writer on an on going basis. Le Compte and Goetz (1982) and Guba (1978) recommend doing member checks of the researcher's data interpretations to test whether or not the conceptual categories are also shared by the participants. Once during the project, and at the end of the project, the teachers verified the typewritten transcripts of data interpretations.

Patterns of similarities and differences between the grade one and the grade two teachers' situations were tested on an on going basis. Although many similarities were found between case studies, how each teacher defined and interpreted her situation was unique to her context, experiences, perceptions, beliefs, assumptions, knowledge, and emotional state of being.

Findings

The two teachers were Franco-Albertans who had received their early education in bilingual schools offering one hour of instruction in French per day. The teachers, who had taught approximately seven years in bilingual and French immersion programs, had not received preservice training for French immersion or for teaching social studies.

The two elementary schools in which the teachers taught fell under the jurisdiction of the same school board. A French immersion program attended by nonfrancophone students and a regular English program were offered in both schools. According to the grade one and grade two teachers, their students came from middle and upper middle class homes. In both schools administrators and librarians were English-speaking. The grade one teacher taught 32 students, and the grade two teacher taught 22 students.

Teachers' Definition of the Situation

The data consistently suggested that the two teachers traded off the social studies concept and inquiry skill development goals prescribed in the *1978 Alberta Social Studies Curriculum* (Alberta Education) for second language development goals. Some of the concepts to be developed at the grade one and two levels include: environment, human needs, identity, institution, and interaction. The major inquiry skills to be developed are: gathering information, analyzing, evaluating and synthesizing data, resolving an issue, formulating a plan of action, and analyzing aspects of human values (pp. 9-24).

The teachers' lack of commitment to the goals of the social studies curriculum appeared to be influenced by their definition of the situation. From a symbolic interactionist perspective, the definition of the situation is a process in which individuals, prior to acting, explore, examine, and feel out through behavior and thought the behavior possibilities of a situation. They will note and evaluate external expectations and demands, including the norms, values, and perceptions of particular groups, and the limitations which the situation imposes upon their prospective line of action (Blumer, 1969, pp. 53-64; Hewitt, 1979, pp. 134-141). As individuals attempt to fit the external elements to their own internal elements such as their intentions, goals, and feelings (Blumer, pp. 53-60), they will form an attitude toward the situation, thereby defining the situation in one way and not in another (Blumer, p. 20; Waller, 1970, pp. 162-163).

Perceptions of Parents and Administrators' Expectations

The two teachers claimed during interviews that parents and school administrators expect that immersion students will acquire French language skills while achieving as well in all areas of schooling as those students in the English program. The teachers considered such expectations as unrealistic.

The grade one teacher explained that all subject matter programs for French immersion are translated from English into French. She said that before she can begin to teach a subject, she has to teach her students the necessary vocabulary. She believes that it is not realistic to expect students with different language learning abilities to have to learn more at a much faster rate than students in the English program.

The grade two teacher stated that parents and administrators expect children

will speak French with ease after spending a year in the French immersion program, and that they will achieve as well, if not better, in all school subjects as those students in the English program. She viewed such expectations as being too much to ask of students who have different learning abilities. In April, she discontinued teaching social studies in order to prepare her students for the school board standardized tests in French and English language arts, mathematics, and science. Social studies, she said, was not tested because it was not considered to be a basic subject in French immersion. She stated that the tests were a way for the school board to verify whether or not immersion students were achieving according to English and French program norms. She thought that the French tests originated from Ottawa. She described school system personnel as lacking an understanding of possible achievable goals in French immersion, and she was worried that if her students did not succeed on the tests, her reputation could be at stake.

Perceptions of Social Studies and Second Language Learning

Either during interviews or after classroom observations, the two teachers showed the researcher an English copy of the *1978 Alberta Social Studies Curriculum* and copies of different school board designed social studies teaching units in French. According to the teachers, it is not possible in an immersion situation to implement the concept and inquiry skill development goals outlined in the provincial social studies document. The teachers' lack of commitment to the goals outlined in the provincial social studies document appeared to be influenced by several factors such as: a lack of training and of human and material resources; their perceptions of consultants' expectations; their perceptions of second language acquisition; and their knowledge, beliefs, and assumptions about teaching, learning, and child socialization.

The lack of resources. The teachers, who had not received any formal training for social studies teaching, stated that without the assistance of competent French-speaking social studies consultants and suitable social studies materials in French, they could not see themselves implementing the provincial social studies concept and inquiry skill development goals.

Consultants' expectations. The teachers explained that the French consultants expect them to teach the vocabulary outlined in the school board designed social studies teaching units, but that it is up to the teachers to select their own methodologies. Even though the teachers described the school board units as being made up of a hodgepodge of unrelated themes (grade one unit topics included the house, the seasons, Thanksgiving, Halloween, Christmas, clothing), and of "little recipes" designed to amuse children, the thematic and vocabulary basis for their classroom social studies curriculum originated mainly from these units. Upon examining the teaching units, the writer noted that these units did not provide the teachers with examples of concept and inquiry skill development activities.

Second language acquisition. The grade one teacher explained that it takes students approximately six to eight months before they can interiorize enough French for their responses to become spontaneous. She described the second language learning process as follows: Firstly, students need to learn to listen so that they can hear the sounds in French. Then they need to repeat the words in French. Before sentences become automatic for them, they go through a "franglais" stage and will say, "Oh! I hurt my jambe!" and "I don't have a gomme to effacer!" The grade one teacher stated that because of her students' limited knowledge of French, it is not

possible for her to involve them in value clarification activities. She said that she does not have the time to ask 32 students questions such as: "How do you feel about such and such?" or "What do you prefer?" She therefore believes that social studies in French immersion can be no more than simple vocabulary development.

The grade two teacher explained that during a discussion, students will either say nothing or use "franglais" or English. They will say a few words in French and expect her to fill in their sentence for them. She was concerned that, when she fills in a sentence for a student, she may not be correctly interpreting what the student has in mind. As far as social studies teaching is concerned, she finds that by the time she teaches her students the vocabulary in the social studies teaching units, she has little time left for concept development and inquiry oriented activities.

Teaching and learning and child socialization. The investigator noted that the two teachers, when planning and teaching social studies, consistently included traditional teacher-controlled vocabulary development and application activities and receptive skill development activities such as: asking students descriptive questions requiring a single correct answer; showing students pictures and asking them to repeat words and sentences; getting students to draw pictures representing new words learned; having students copy words in their scribblers; involving the class in singing songs containing the vocabulary taught; and reading to students stories containing newly introduced vocabulary.

The two teachers during interviews described themselves as including citizenship development goals in their social studies teaching. They interpreted citizenship development as child socialization and life skills development. The writer noted that, during classroom observations, both teachers generally encouraged their students to share, to be patient and tolerant towards others, and to obey classroom rules.

Discussion and Implications

Influenced by a complex interplay of external and internal elements, the teachers, when planning and teaching social studies, traded off concept and inquiry skill development goals for second language development goals, and they consistently involved their students in teacher-controlled vocabulary and receptive skill development activities. One could argue that if social studies in French immersion is taught using traditional second language teaching methodologies, then immersion students' opportunities to develop important citizenship goals and to practice inquiry skills such as information seeking, analyzing and evaluating information and values in conflict, drawing conclusions, making inferences, developing concepts and generalizations, and solving personal and social problems, may be limited.

The Need for a Commitment to the Goals of Social Studies

The two teachers' decisions to trade off social studies concept and inquiry skill development goals for second language development goals, especially vocabulary development, and the development of receptive skills, may be a normative practice in other immersion contexts.

There is evidence which seems to suggest that the importance of language development is being overemphasized in French immersion (Shapson, 1984). Wightman (1979), for example, has observed that in early grades, the immersion teacher's initial step when introducing a new topic for study is to develop students'

vocabulary (p. 59). There is further evidence which suggests that receptive skill development, rather than productive skill development, is a primary objective in French immersion. For example, Lapkin (1984) has noted that for years the primary objective in the early stages of any French immersion program was to ensure the development of students' receptive skills (p. 575). Harley (1984), who conducted an in-depth study of elementary immersion students' grammar, discourse, and sociolinguistic competencies has found that immersion students have a remarkable ability to comprehend spoken and written discourse. She attributes these abilities to the methodology actually used in French immersion, such as students listening to the teacher talk and reading subject matter texts.

Recently, Long and Porter (1985) have criticized the "comprehensible input" approach to second language acquisition, and have questioned whether simply hearing and understanding the new items are sufficient for a learner to use the language. They call for the need to give students opportunities to produce new forms of language in group work (p. 214). Likewise, Allen (1985) sees the need to use language for real purposes in a student-centered class which gives students "a chance to communicate their thoughts and feelings in a second language" (p. 996). Allen also recommends asking students "evaluative" questions rather than factual questions, as evaluative questions give them a chance to communicate their thoughts and feelings, without having to worry about errors in grammar or pronounciation and can result in a higher level of communication (pp. 995-996). According to Long and Porter (1985), even if students make errors, these should be viewed as a "healthy" part of language development (p. 223). Harley (1984) has found that immersion students are capable of responding in French in a manner that is usually coherent in the context, and that they will make use of nonverbal communication strategies such as gesture or mime to get their meaning across. She has noted that without having mastered conditional verb forms (verb forms necessary for expressing a hypothetical situation in social studies), grade five students will use a variety of verb forms such as the present tense together with an adverb, or the future construction with an infinitive, to get the essential meaning across.

Calling upon second language teachers to shift from a teacher-centered, receptive skill development approach, to a student-centered, productive skill development approach may be problematic. There is considerable evidence which suggests that teacher talk and receptive skills are also being emphasized in English classrooms (Ponder & Davis, Jr., 1982; Shaver, Davis, & Helburn, 1979; Sirotnik, 1983; Stodolsky, 1984; VanSickle, 1985). Resistance to the new social studies is also a case in point. Expository teaching, primarily via textbook and teacher talk, continues to dominate practice in most of today's social studies classrooms (Engle, 1986, p. 20). Engle points out that in most schools, the intellectual and social climate and governance are "unfavorable to open, questioning education over which the learner exercises some measure of control" (p. 22). Sirotnik (1983) explains that the greatest difficulty facing teachers today who may want to adopt a more student-centered or inquiry approach to instruction, is that they must overturn an enduring stereotypical image of what is a "good" classroom, or what Enright and McCloskey (1985) refer to as the "No talking!" image. Teachers, Engle (1986) argues, are the products of a "mindless education," and they are "poorly prepared to help students learn the skills of social criticism that are so important to democracy" (pp. 21-22).

Preservice and inservice teacher education. If teachers and future teachers of French immersion are to develop a commitment to the purpose and goals of social

studies, there is a need during their preservice and inservice training to: (a) allow them to inquire into the problems and questions of various social science disciplines; (b) help them to understand that the goals of immersion and of social studies programs go far beyond the mastery of vocabulary and grammatical structures, and that language learning is not an end in itself, but rather a tool for learning something else; (c) make them aware that social studies concept and inquiry development approaches have many similarities with the communicative competence approach, in that both approaches are learner-centered and focus on helping students develop meaning, see the relationships involved in the materials studied, and develop their inferencing and problem solving abilities; and (d) give them opportunities to practice in a laboratory setting (Berliner, 1985) communicative competence and inquiry skill development techniques such as simulations, group work, role playing, and questioning strategies for problem solving and inferencing.

The need for a realistic social studies program. The finding from the case study, that French immersion students are expected by parents and administrators to achieve as well as students in the English program, may be a normative expectation in other immersion milieus. The literature on programs of studies for French immersion across Canada (Conseil des Ministres de l'Education, Canada, 1983; Lapkin, 1984) confirms that immersion students are expected to study the same content in all subject areas as their peers in the regular English program, and that subject matter programs are generally translated from English into French.

Integrating communicative competence goals and social studies goals. There appears to be a need to develop a realistic social studies program for early elementary French immersion which goes beyond the translation of English documents into French. Such a program could incorporate examples of social studies concept and inquiry development activities integrated with communicative competence techniques suggested by second language researchers, notably Enright and McCloskey (1985), Long and Porter (1985), Maiguashca (1984), and Yalden (1984). A more controlled approach to concept development would have to be considered in such a program. The vocabulary for dealing with the attributes of the concepts would probably have to be specified. As well, examples of grammatical structures could be included which would allow students to respond to analytical and evaluative types of question.

Research and debate. Unless the question of trade-offs in French immersion is further explored and openly debated, what appears to be a successful innovative approach to second language learning could limit students' opportunities to develop critical thought. There is, therefore, a need for research which could provide insights into how immersion teachers' beliefs, assumptions, and their second language teaching strategies can either enhance or limit students' opportunities to develop social studies concepts and to acquire the knowledge, skills, and values needed for democratic citizenship.

References

Alberta Education. (1978). *1978 Alberta social studies curriculum* (Interim ed.). Edmonton: Alberta Education.

Allen, E.D. (1985). Communicative competence and levels of proficiency. *The Canadian Modern Language Review, 41*(6), 991-999.

Barik, H.C., & Swain, M. (1976). Primary-grade French immersion in a unilingual English-Canadian setting: The Toronto study through grade two. *Canadian Journal of Education, 1*(1), 39-58.

Berliner, D.C. (1985, November-December). Laboratory settings and the study of teacher education. *Journal of Teacher Education*, 2-8.

Bibeau, G. (1984). No easy road to bilingualism. *Language and Society, 12*, 44-47.

Black, J., & Champion, D.J. (1976). *Methods and issues in social research*. New York: John Wiley.

Blumer, H. (1969). *Symbolic interactionism. Perspective and method*. Englewood Cliffs, NJ: Prentice-Hall.

Calgary Catholic Schools. (1979). *Report on Calgary Catholic Schools French immersion program, E.C.S. and grade one, 1978-1979. A research review and comparative evaluation.* Calgary, Alberta.

Conseil des Ministres de l'Education (Canada). (1983). Actualités provinciales. *Dialogue, 1*(4).

Council of Ministers of Education (Canada). (1982). Manitoba. The French immersion research project. *Dialogue, 1*(2).

Cummins, J. (1976). The influence of bilingualism on cognitive growth. A synthesis of research findings and explanatory hypotheses. In M. Swain & G. Dumas (Eds.), *Working papers on bilingualism, 9*, (pp. 1-43). Toronto: The Ontario Institute for Studies in Education.

Cummins, J. (1983). Research findings from French immersion programs across Canada: A parent's guide. *CPF Special Report.* Ottawa: Canadian Parents for French.

Day, E.M., & Shapson, S. (1985). Assessing oral communicative skills in immersion programs. *CAIT News, 8*(3), 15-17. Ottawa: Canadian Association of Immersion Teachers.

Delamont, S. (1976). *Interaction in the classroom*. London: Methuen.

Denzin, N.K. (1970). The methodologies of symbolic interaction: A critical review of research techniques. In G.P. Stone & H.A. Farberman (Eds.), *Social psychology through symbolic interaction* (pp. 447-465). MA:Ginn.

Engle, S. (1986). Late night thoughts about the new social studies. *Social Education, 50*(1), 20-22.

Enright, D.S., & McCloskey, M.L. (1985). Yes, talking! Organizing the classroom to promote second language acquisition. *TESOL Quarterly, 19*(3), 431-453.

Esland, G.M. (1971). Teaching and learning as the organization of knowledge. In M.F.D. Young (Ed.), *Knowledge and control. New directions for the sociology of education* (pp. 70-116). London: Collier-Macmillan.

Genesee, F. (1984). French immersion programs. In S. Shapson & V. D'Oyley (Eds.), *Bilingual and multicultural education: Canadian perspective* (pp. 33-34). Avon, England: Multilingual Matters.

Gray, V.A., & Cameron, C.A. (1975). *Evaluation of early French immersion in Fredericton, NB* (Interim Report). Fredericton: Psychology Department, University of New Brunswick (mimeo).

Guba, E.G. (1978). *Criteria for assessing the trustworthiness of naturalistic inquiries* (ERIC/ECTJ Annual Review Paper, 29, 75-91). Syracuse, NY: Syracuse University.

Harley, B. (1984). How good is their French? *Language and Society, 12*, 55-60.

Helmstadter, G.C. (1970). *Research concepts in human behavior*. New York: Appleton-Century Crofts.

Hewitt, J.P. (1979). *Self and society: A symbolic interactionist social psychology* (2nd ed.). Toronto: Allyn and Bacon.

Krashen, S.D. (1984). Immersion: Why it works and what it has taught us. *Language and Society, 12*, 61-64.

Laberge, E. (1984). The French immersion programme in St. Norbert. *Contact, 3*(4), 9-11.

Lapkin, S. (1984). How well do immersion students speak and write French? *The Canadian Modern Language Review, 40*(4), 575-585.

Lapkin, S., & Swain, M. (1984). Research update. *Language and Society, 12*, 48-53.

LeCompte, M.D., & Goetz, J.P. (1982). Problems of reliability and validity in ethnographic research. *Review of Educational Research, 52*, 31-60.

Li, P. (1981). *Social research methods*. Toronto: Butterworth.

Long, M.H., & Porter, P.A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207-228.

Mahé, Y.T.M. (1984). *Social studies teaching and French immersion: A case study of two elementary teachers' perspectives*. Unpublished doctoral dissertation, University of Alberta.

Maiguashca, R.U. (1984). Semantic fields: Towards a methodology for teaching vocabulary in the second-language classroom. *The Canadian Modern Language Review, 40*(2), 274-297.

Mead, G.H. (1934). *Mind, self and society*. (C.W. Morris, Ed.). Chicago: University of Chicago Press.

Nelson, M., & Cornbleth, C. (1982). Social studies research and classroom teachers: Open the door wider. *Social Education, 46*, 218-220.

Pohland, P. (1972). Participant observation as a research methodology. *Studies in Art Education 13*.

Ponder, G., & Davis, O.L. Jr. (1982). Social studies education. In H.E. Mitzel (Ed.), *Encyclopedia of educational research* (5th ed.) (*4*, pp. 1723-1732). New York: American Educational Research Association.

Shapson, S.M. (1984). Personal comments on immersion in Canada. *Contact, 3*(1), 2-3.

Shapson, S., & Kaufman, D. (1976). French immersion: A western perspective. *Canadian Society for the Study of Education, Year Book, 3*, 8-25.

Shapson, S., & Kaufman, D. (1978). Overview of elementary French programs in British Columbia: Issues and research. *The Canadian Modern Language Review, 34*(3), 586-603.

Shavelson, R.J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*, 455-498.

Shaver, J.P., Davis, O.L., & Helburn, S.M. (1979). An interpretive report on the status of precollege social studies education based on three NSF funded studies. *Social Education, 43*, 3-18.

Sirotnik, K. (1983). What you see is what you get: Consistency, persistency, and mediocrity in classrooms. *Harvard Education Review, 53*(1).

Smith, L.M., & Geoffrey, W. (1968). *The complexities of an urban classroom: An analysis toward a general theory of education*. New York: Holt, Rinehart & Winston.

Stern, H.H. (1984). The immersion phenomenon. *Language and Society, 12*, 4-8.

Stodolsky, S.S. (1984). Teacher evaluation. The limits of looking. *Educational Researcher, 13*(9), 11-18.

Swain, M., & Lapkin, S. (1981). *Bilingual education in Ontario: A decade of research*. Toronto: Ontario Institute for Studies in Education.

Swain, M., & Lapkin, S. (1983). *Evaluating bilingual education: A Canadian case study*. Avon, England: Multilingual Matters.

VanSickle, R.L. (1985). Research implications of a theoretical analysis of John Dewey's How We Think. *Theory and Research in Social Education, 13*(3), 1-20.

Waller, W. (1970). The definition of the situation. In G.P. Stone & H.A. Farberman (Eds.), *Social psychology through symbolic interaction* (pp. 162-174). Waltham, MA: Ginn.

Wesche, M.B. (1984). A promising experiment at Ottawa University. *Language and Society, 12*, 20-25.

Wightman, P. (1979). French immersion: An educational alternative. In B. Mlacak & E. Isabelle (Eds.), *So you want your child to learn French! A handbook for parents* (pp. 54-65). Ottawa: Canadian Parents for French.

Wilson, S. (1977). The use of ethnographic techniques in educational research. *Review of Education Research, 47*, 245-265.

Woods, P. (1985). Sociology, Ethnography and teacher practice. *Teaching and Teacher Education, 1*(1), 51-62.

Yalden, J. (1984). The design process in communicative language teaching. *The Canadian Modern Language Review, 40*(3), 398-413.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 212-222.*

WALTER B. BOLDT

and

BILLIE HOUSEGO
University of British Columbia

Critical Incidents in the Supervision of Student Teaching: A Cluster Analysis of the Perceptions of Graduate Transfer Students

To gain further understanding of student teachers' perceptions of their experiences in having their teaching supervised, student teachers were asked to rate a sample of critical incidents in the supervision of student teaching on three, seven-point, bipolar scales anchored by the terms realistic-unrealistic, important-unimportant, and manageable-unmanageable. They were asked three questions concerning these incidents: (a) To what extent do these incidents reflect the reality of the supervision experience? (How common are these incidents in the supervision of student teaching?) (b) To what extent are these incidents critical or important in the supervision of student teaching? (Would discussing these incidents with your supervisors affect your teaching performance?) (c) To what degree are the problems described in the incidents ones on which student teachers can take some action; how manageable are they? (Could you respond in some meaningful way to such issues?) The rating profiles of the incidents over the three scales were compared through hierarchical cluster analysis. Three clusters, personal adjustment incidents, lesson preparation incidents, and housekeeping rules of conduct incidents along with indications of their perceived realism, importance, and controllability emerged. The results confirm the view that "survival type" incidents were central in the student teachers' supervision experiences and that supervisors of student teaching are, for the most part, ideological in their supervisory approach.

The all-encompassing goal of supervising student teaching is to ensure that the student teacher becomes proficient in teaching. This assumes that a supervisor, whether a classroom teacher or a university professor, can identify effective

---

Dr. Boldt, whose interests are in research design and measurement, is a member of the Faculty of Education.

Dr. Housego is also a member of the Faculty of Education and has interests in teacher preparation and the supervision of student teaching.

teaching and guide the student teacher toward that goal, that he or she can analyze the discrepancy between "what is" and "what ought to be" and meet the student teacher's need for "a sustained, expert program to help him/her relinquish . . . existing classroom behavior in favor of new behavior" (Cogan, 1973, p. 4). More fundamentally, it also assumes that supervisors are able to perceive teaching in the manner that student teachers perceive it, and that they can accommodate their supervisory behavior to these perceptions in order to bring about desired teaching behaviors. Basic questions about what supervisors actually do in an effort to be helpful must be asked if the supervisory process is to be improved; moreover, it is important to know how student teachers perceive the help provided them. Which actions on the part of supervisors are perceived to contribute to improved teaching and which are viewed as less helpful? Supervisors of student teaching vary considerably in the approaches they use in assisting and providing feedback to their supervisees. A number of such approaches are summarized in a model of student teaching supervision described by Housego and Grimmett (1983).

The categorization of supervisory behavior proposed by Housego and Grimmett, and shown in Figure 1, is rooted in the supervisor's need to take account of two major concerns, to reduce the threat often felt by student teachers and to make the best possible use of the limited time available in brief periods of school experience. As shown on the left-hand side of Figure 1, if threat concerns are uppermost, a supervisor attempts to be facilitative or understanding and assuring. If the supervisor utilizes a research-based knowledge of teaching, supervision can go beyond being purely indulgent, a so-called "tea party" (Goldhammer, 1969, p. 61), to become to a greater degree knowledge-based or deliberative. The supervisor functioning in the deliberative mode "is able to discuss with the student teacher alternative strategies for achieving desired outcomes with pupils—outcomes contingent upon learner characteristics and the nature of the classroom environment" (Housego and Boldt, 1984, p. 170). The deliberative supervisor is knowledgeable with respect to effective teaching, whereas the indulgent supervisor is simply a kind, concerned person.

When shortage of time concerns are uppermost, the supervisor may assume a directive role, as shown on the right-hand side of Figure 1, and, depending on the degree to which supervision is knowledge-based, either impose strong views about teaching on the student teacher in an ideological mode of supervision or, with a greater knowledge of teaching, proceed in a "technological" mode in which "teaching is regarded as being composed of a definable repertoire of knowledge, skills, and attitudes that a teacher brings to bear in an effort to create certain changes in learners" (Zumwalt, 1982, pp. 223-224). The technological supervisor, like the deliberative supervisor, is knowledgeable with respect to effective teaching, whereas the ideological supervisor may know from experience "what works" and insist that the student teacher adopt those approaches.

Assuming either a predominantly facilitative supervisory role or a predominantly performance-based one is likely to create an undesirable polarity. If neither position rests in research-based knowledge of teaching, the resulting entrenched either ideological or indulgent supervision modes readily degenerate to become impoverished supervision as shown at the bottom of Figure 1. What is needed is greater flexibility on the part of supervisors, the adoption of a so-called metadeliberative position (see Figure 1) in which "throughout the process of practicum supervision . . . supervisors constantly deliberate . . . on which of two supervisory modes most appropriately meets the apparent and most pressing needs

of the specific practicum context in which they must supervise at a given time" (Housego and Grimmett, 1983, p. 333).

The dual concerns of time and threat are thus taken into account—technological supervisors going beyond merely presenting research proven skills and having student teachers practice them, to also concentrate as required, on students' needs to feel secure and confident in the school environment, and deliberative supervisors extending their "research-embracing" facilitative approach to engage student teachers in practicing and receiving feedback on specific instructional behaviors. Emphasis may shift and at times be more towards pupil learning and instructional performance and yet, at other times, move more toward developing the confidence and professional autonomy of the student teacher.

Theoretically, a metadeliberative approach to supervising student teaching has great appeal; however, supervisors' abilities to function in this mode depend on their being able to understand student teachers' perceptions of their classroom and supervisory experiences—the students' reality.

In an effort to explore this reality, a study of critical incidents occurring in the supervision of student teaching was undertaken (Housego & Boldt, 1984). One hundred, twenty-five critical incidents in the supervision of student teaching were collected from elementary and secondary student teachers with varying backgrounds and experience in education, from supervision teachers, and from faculty supervisors. Because it was expected that incidents in student teaching supervision would stem from the total spectrum of problems encountered in teaching, the incidents were categorized according to Broudy's (1965) categories of problems of practice which characterize the education profession: aims and objectives, organization-management, curriculum, and teaching-learning. It was found that nearly 50% of the supervision incidents grew out of concerns with organizing and managing instruction. In another 30% of the supervisory incidents identified as critical, interpersonal relationships between student teachers and their supervisors were predominant. This necessitated the formation of an additional, so-designated category beyond Broudy's classifications. Few supervisory incidents, about 16%, were rooted in substantive problems in teaching and learning. Critical supervisory incidents having to do with educational goals or curriculum-based problems were conspicious by their absence. These findings generated additional questions, particularly with respect to the relative importance of the incidents, the degree to which they are truly representative of the experiences of student teachers generally, and the degree to which students perceive that they can manage or take action to remedy the problems described in some of the incidents. Accordingly, critical incidents contributed by student teachers were isolated and a decision made to ask another somewhat homogeneous group of student teachers to respond to such questions in a second phase of the study, an account of which is presented in this article.

Method

Subjects

It was expected that subjects would respond to critical incidents according to their experience and background in education. To begin mapping these responses, 45 student teachers from the one-year post-baccalaureate program of the Faculty of Education at the University of British Columbia were selected to provide the data to be studied. They constituted about 20% of the student teachers in this program.

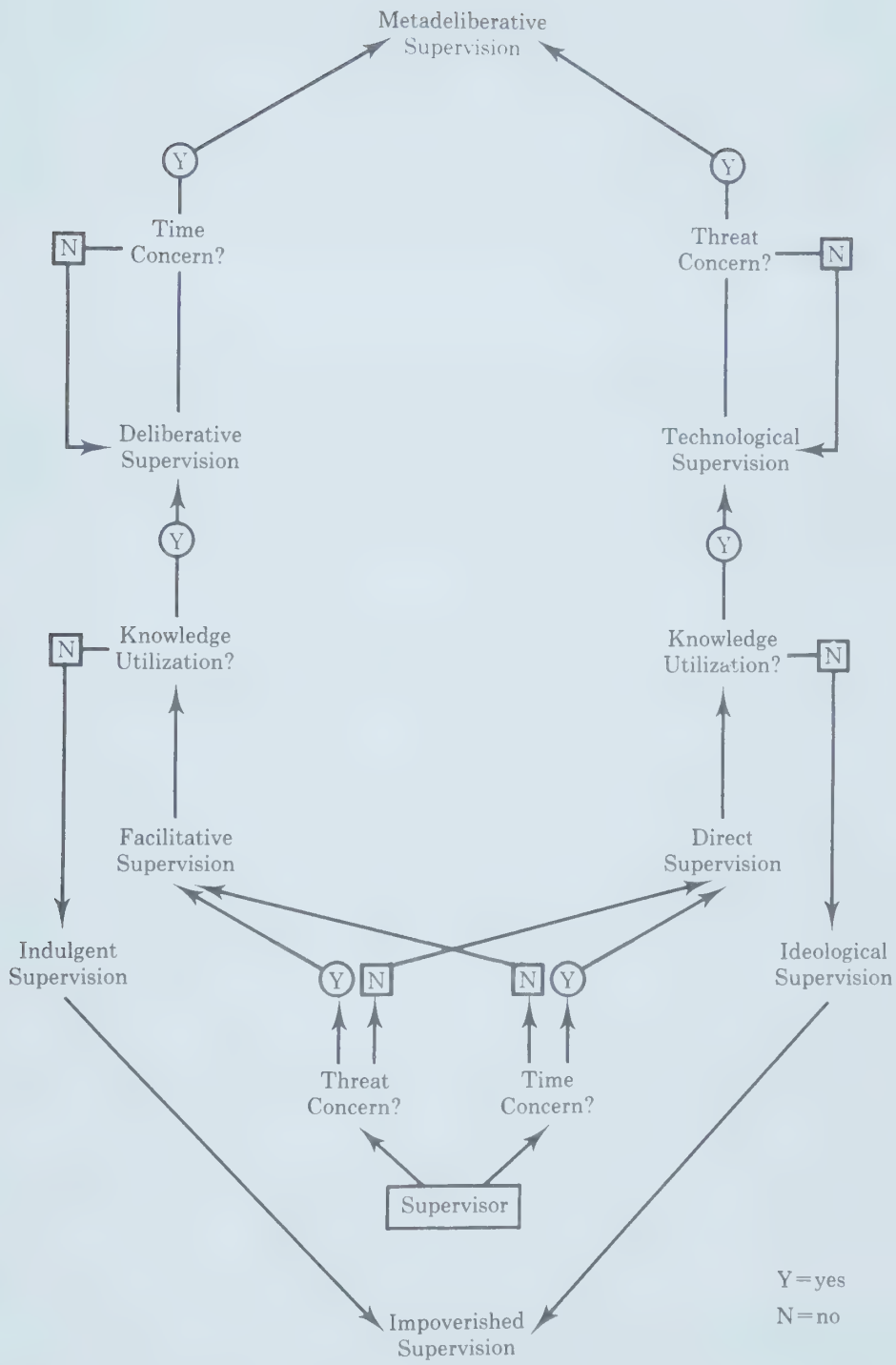

Figure 1: A typology of approaches and the metadeliberative perspective on supervision.

The educational experiences of these subjects included a degree from another faculty, and, at the point where the data were collected, five months of professional studies ranging from philosophy of education to methods of teaching various school subjects, plus seven weeks of school experience. In this one-year program, the student teachers are relatively homogeneous in terms of background and experience in education and, upon successful completion of the program, are eligible for certification as teachers. These were considerations in selecting the subjects in order to ensure reasonable uniformity and comprehensiveness in the background and experiential knowledge on which their responses to the critical incidents would be based. The 45 subjects were also drawn in such a way as to ensure reasonable representation of elementary and secondary preservice teachers, the latter group including a cross section of subject matter specialists.

Items

Since not all 125 incidents collected in the first phase of the study could be used in the second phase, and in order to avoid any systematic bias in selecting a feasible number of items, a random sample of 30 incidents was drawn from the pool of student-generated incidents. The sample size of 30 was determined on the basis of practical considerations such as time and the fatigue of the respondents. Trials indicated that about 30 items would be manageable.

Instrument

From the standpoint of using the results of this study for the improvement of student teaching supervision, particularly in the one year program, it seemed important to determine the degree to which subjects would see the incidents: first, as realistic, second, as important in terms of their own experience in being supervised, and third, to the extent that incidents reflected problems, as manageable under the constraints of the student teaching experience. The instrument devised to obtain this information consisted of three 7-point, bipolar adjectival scales anchored by the terms realistic-unrealistic, important-unimportant, and manageable-unmanageable applied to each of the 30 items as shown in the following sample item.

When pupils argued and failed to help set up P.E. equipment, my faculty advisor gave me explicit directions as to how to handle setting up equipment efficiently.

REALISTIC: ..... : UNREALISTIC
7 6 5 4 3 2 1

IMPORTANT: ..... : UNIMPORTANT
7 6 5 4 3 2 1

MANAGEABLE: .....: UNMANAGEABLE
7 6 5 4 3 2 1

Analysis

An analysis of the ratings for each of the 30 incidents on each of the three scales was carried out in order to determine whether or not the scales would serve to differentiate among the incidents and, in addition, if the three scales each would differentiate among the incidents in a unique way. The method of analysis used

was a repeated measures, analysis of variance design, with one within-subjects factor (scales) and one between-subjects factor (incidents). The number of observations in each of the 90 cells was 45. The three scales, realism, importance, and control constituted the three levels of the within-subjects factor and served as bases of judgment for differentiating between the incidents. The scales also allowed the subjects to indicate the degree to which they perceived each basis of judgment (scale) to apply to the incidents. By this method, then, the way student teachers perceived critical incidents in the supervision of student teaching from the standpoint of how realistic the incidents were, how important the incidents were in their experience, and how manageable they considered the incidents to be from their perspective, were obtained.

Results

Repeated-Measures Analysis

Table 1 gives the results of the analysis of the ratings of the incidents on the three scales. It shows a significant difference among scales, a significant difference among incidents, and significant interaction effects between scales and incidents. These results were interpreted as indicating that, in terms of the bases of judgments represented by the scales, the student teachers did perceive the incidents differentially and, hence, provided at least some insight into what topics or areas in the supervision process they experienced that they perceived to be realistic, important, and manageable. For supervisors, then, it would seem advisable to accommodate these perceptions in planning a metadeliberative approach.

TABLE 1

Analysis of Variance Based on Repeated Measures

Source of Variation	Sum of Squares	Df	Mean Square	F
A	833.00	29	28.76	7.17*
S-Within	5289.31	1319	4.01	
B	959.04	2	479.52	270.72*
AB	283.62	58	4.89	2.76*
BS-Within	4672.56	2638	1.77	

* $p<0.01$

Hierarchical Cluster Analysis

The hierarchical clustering analysis compares the rating profiles of each incident over the three scales, obtains a measure of the similarity of the profiles of each incident over the three scales, obtains a measure of the similarity of the profiles, and progressively groups the incidents in such a way that the rating profiles are as similar as possible and the variation of the ratings in each group formed is as small as possible (Ward, 1963). At each step in the grouping process, the within-group variation (error) is minimally increased. A substantial increase in the error at a particular step suggests the beginning of a new cluster or group.

Table 2 reproduces the essential groupings displayed in the tree-diagram, the identification number of the incidents, and the mean scores of the incidents on each of the three scales.

Interpretation of the clusters of incidents in Table 2 was based on a close examination of the interrelationships of the substantive content of the incidents as well as the mean ratings on the scales. Beginning with the innermost cluster in each of the three major categories (A, B, and C) and working outwards, the following interpretation of the three groupings of clusters of critical incidents in the supervision of student teaching was obtained:

Category A (Supervision Incidents Emanating from Personal Adjustment Problems)

Many of the problems of practice central to the incidents in Category A are what Brophy (1983, p. 282) would refer to as "personal adjustment problems." In particular, disruptive self-criticism and coping with unusually disruptive behavior in individual pupils are the focal problems. Other clusters in this category included personal adjustment to somewhat less disruptive behaviors of the class as a whole (noise, inattention) and of oneself (poor planning habits). All the incidents in this category were rated as realistic, important, and manageable. These results suggest that coping personally with unusually disruptive behaviors on the part of oneself and, especially, the class is still very much a part of the experience of the subjects. Moreover, as Tardif (1985) observes, managing to control one's class is an integral part of one's "teacher identity." As Fuller and Brown (1975, p. 37) might view them, the supervision incidents in Category A appear to relate largely to teaching problems which could be categorized as "survival concerns," representative of the first stage in learning to teach. The supervisory approach, in most instances, was to supply some fairly precise suggestions or directives as to what to do. Seldom do the student teachers' critical incident accounts of the responses of their supervisors reveal that alternative strategies or research-based knowledge were discussed. The supervisory approach could be seen as largely ideological, although a more facilitative approach might have been expected.

Category B (Supervision Incidents Emanating from Lesson Presentation Problems)

This category shows a trend away from basic survival concerns toward concerns about teaching a "good" lesson. Specifically, the supervision incidents relate to problems of presenting an unappealing lesson, teaching difficult concepts, keeping the attention of the class, dealing with talkative students during the lesson presentation, and concluding the lesson successfully. The subjects rated these supervision incidents as realistic, very important, and quite manageable. Overall, the incidents in this category received the highest ratings. While these incidents were still closely tied to survival concerns, they do suggest a transition toward a somewhat more advanced focus in the supervision of student teaching, a stage in which the problems have to do with group management during active instruction and with lesson presentation itself. One would expect that such problems would most often be approached in a somewhat more direct and technological mode of supervision (Housego & Grimmett, 1983, p. 324). Although the supervisor's response in dealing with these problems was generally direct, there is no evidence in the student teachers' critical incident statements of their supervisor's responses that learnings

TABLE 2

Mean Ratings and Clustering of Incidents

Major Categories of Clusters (30 Incidents)		Mean Ratings* (N=45) Realism	Importance	Control
A	2	5.2	6.2	5.7
	25	5.3	6.2	5.6
	26	5.3	6.2	5.6
	29	5.1	6.3	5.5
	9	5.4	6.0	5.4
	20	4.9	5.7	5.8
	23	4.8	5.8	5.8
	4	5.1	5.9	5.9
	6	4.3	5.9	5.4
	30	4.4	5.9	5.6
	28	4.6	5.6	5.5
	10	4.4	6.1	5.8
	19	4.5	6.1	5.9
B.	11	5.2	6.0	6.1
	18	5.1	6.1	6.3
	7	5.1	6.3	6.3
	3	5.5	6.3	6.2
	21	5.3	6.2	6.0
	13	5.6	6.6	6.1
	27	5.8	6.6	6.2
C.	12	4.3	5.2	5.3
	14	4.5	5.2	5.4
	1	4.4	5.2	5.2
	5	4.6	5.2	5.9
	15	4.1	5.7	5.2
	22	4.0	5.9	5.2
	24	3.6	4.9	5.1
	8	3.5	5.5	4.3
	16	3.3	5.6	4.9
	17	4.2	6.5	4.5

* 7-point, bipolar adjectival rating scales

based on research were drawn to their attention. The approach was again ideological, primarily offering "craft-rules" for dealing with problems.

Category C (Supervision Incidents Emanating from Housekeeping and Rules of Conduct Problems)

The supervision incidents in Category C can be characterized according to Brophy (1983, p. 282) as relating to the "development of a workable set of housekeeping procedures and conduct rules." The characteristic incidents deal with preparing and distributing materials and equipment, handling incidents, checking homework, using the blackboard efficiently, getting the attention of the class, evaluating achievement, and planning a suitable lesson. Based on the ratings, these supervision incidents are seen as only marginally realistic, somewhat important, and manageable. Supervisory incidents in this category, too, are likely to be approached in a technological mode, lending themselves as they do to specific prescriptions for action. Supervisors appear to have been quite directive, relying heavily on craft-rules. There is little evidence in what the student teachers report of supervisor reaction, that supervisors use research-based knowledge of teaching to engage student teachers in discussions of alternative problem solving strategies.

In summary, the results suggest, that despite the category of incidents under consideration, according to the student teachers' critical incident reports, supervisors tended to be largely ideological in their approach, using craft-rules more than research-based knowledge of teaching, and, therefore, being unlikely to provide student teachers with the intellectual maneuverability which would increase their ability to consider and select from a variety of problem solutions.

Conclusions

Critical incidents in the supervision of student teaching can be identified by any member of the student teaching triad: the student teacher, the university supervisor, or the sponsor teacher. The major thrust of this study was to examine more closely a sample of critical incidents in the supervision of student teaching generated by student teachers. In addition to the usual method of categorizing procedures typical of critical incident analysis (Housego & Boldt, 1984), the writers in this study sought to enhance the analysis through the use of an empirical technique, cluster analysis, to identify focal points for more intensive study. The cluster analysis did group the incidents in an interesting and interpretable way. Specifically, the analysis produced a differentiation among these incidents which agreed with that postulated by Fuller and Bown (1975) at the first stage of learning to teach, the stage of "survival concerns," and with the centrality of the problems of classroom organization and management described by Brophy (1983). The cluster analysis of the incidents also provided a useful adjunct to the usual mode of research in critical incident studies, in providing explicit empirical information on how the clusters were developed and how the incidents in each cluster were interrelated.

On the substantive side of the study, as opposed to the methodological, the results draw attention to needed training in both the craft aspects of professional activity and the application of pedagogical knowledge. Brophy (p. 266) notes that, in practice, the two key teaching tasks, classroom management (craft aspects of teaching) and instruction (pedagogy) are interdependent. This interdependence is clearly evident in the tree-diagram of the clusters making up the three major

categories shown in Table 2. In addition, Brophy contends that,

> Clear and detailed information is now available about how successful teachers organize and manage their classrooms . . . if applied systematically, these principles will enable teachers to establish their classrooms as effective learning environments and to cope with most conduct problems. (p. 266)

The cluster analysis shows that student teachers attach substantial importance to such incidents and that they are viewed as very characteristic of their school and related supervision experiences. This is of relevance to supervisors of student teaching, suggesting the applicability of an approach to supervision such as that detailed in Figure 1, in which both the technology (craft-like aspects of teaching based on research) and concerns with the development of the student teacher's confidence and autonomy come together according to the immediate needs of the student in a particular situation.

Student teachers' perceptions of supervision experience require further study. They must be taken into account by effective supervisors and included in meaningful supervisory training programs. Considering the evidence provided by this study, one becomes aware of the desirability of supervisors, whether classroom teachers or university faculty, drawing on the presently available research-based knowledge of teaching to enable themselves to assume a metadeliberative approach and thereby enhance the integration of the school experience and course work components of teacher preparation. In our view, the supervisory process includes dealing with many of the factors to which Campbell (1985) refers in her study of the variables affecting the instruction behavior of physical education student teachers during their practicum.

References

Brophy, J. (1983). Classroom organization and management. *The Elementary School Journal, 83*(1), 265-285.

Broudy, H.S. (1965). Criteria for the professional preparation of teachers. *Journal of Teacher Education, 16*(4), 408-415.

Campbell, C.E. (1985). Variables which influence student teacher behavior: Implications for teacher education. *The Alberta Journal of Educational Research, 31*(4), 258-269.

Cogan, M.L. (1973). *Clinical supervision.* Boston: Houghton-Mifflin.

Fuller, F.F., & Bown, O.H. (1975). Becoming a teacher. In K. Ryan (Ed.), *Teacher Education* (pp. 25-52). Chicago: National Society for the Study of Education.

Goldhammer, R. (1969). *Clinical supervision.* New York: Holt, Rinehart & Winston.

Housego, B.E.J., & Boldt, W.B. (1984). Critical incidents in the supervision of student-teaching. In P. Grimmett (Ed.), *Research in teacher education: Current problems and future prospects in Canada* (pp. 167-192). Vancouver: CSTE/CSCI/CATE, Faculty of Education, University of British Columbia.

Housego, B.E.J., & Grimmett, P.P. (1983). The performance-based/developmental debate about student teaching supervision: A typology and a tentative resolution. *The Alberta Journal of Educational Research, 29*(4), 319-337.

Tardif, C. (1985). On becoming a teacher: The student teachers' perspective. *The Alberta Journal of Educational Research, 31*(2), 139-148.

Ward, J.H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association, 58*(301), 236-244.

Zumwalt, K.K. (1982). Research on teaching: Policy implications for teacher education. In A. Lieberman and M.W. McLaughlin (Eds.), *Policy making in education* (pp. 215-248). Chicago: National Society for the Study of Education.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 223-233.*

JODI GRANT
University of Saskatchewan

and

CAROL BROWN
Saskatoon Public Schools

Precocious Readers: A Comparative Study of Early Reading and Non Early Reading Kindergarten Entrants

This study sought to determine what percentage of first time English-speaking kindergarten entrants could be identified as early readers (ERS), and to identify differences between ERS and nonearly readers (NERS) with respect to their home background, home learning experiences, and book orientation skills. In September 1984, all first-time kindergarten entrants in the Saskatoon Public School System were screened using the Durkin Word Test *(1966) and the* Durrell Analysis of Reading Difficulty *(1980). Twenty students identified as ERS were randomly matched with NERS from their classrooms. All students were then administered the* Book Orientation Scale. *Additional information was obtained through a parent questionnaire and interviews. Analysis of the data led to the following conclusions. ERS represented 1.64% of the population of kindergarten entrants. Demographic differences between ERS and NERS were similar to those found in previous research studies. The home environments of ERS appeared to have been conducive to the development of skill in reading and to the development of strategies for learning in a wide range of areas.*

The process by which children acquire literacy skills at an early age has been, and continues to be, the focus of much discussion and research. In 1908, Huey described the "natural reader" or the preschool child who learns to read at home without having participated in formal reading instruction. In more recent years, a number of comparative studies have attempted to establish why some children enter kindergarten already reading (Carlson, 1975; Clark, 1976; Durkin, 1966;

Dr. Grant is an Assistant Professor in the Curriculum Studies Department currently involved in research on multiculturalism and folk literature.

Ms. Brown is a resource room teacher and recently completed her Master of Education degree at the University of Saskatchewan.

Evans & Smith, 1976; Getman, 1979; Harty, 1975; King & Friesen, 1972; Manning & Manning, 1984).

Logically, a knowledge of the number of kindergarten entrants who can be identified as early readers is important for long- and short-term planning on the part of school jurisdictions. The literature raises some questions as what percentage of kindergarten entrants can be expected to be early readers. In 1966 Durkin found less than one percent of the population could be identified as early readers, while in 1972 King and Friesen identified just over one percent of the population as early readers. Harty (1975) found six percent of the population to be reading at the time of school entrance. Three years after Harty's study, Durkin (1978) stated that more and more preschoolers were learning to read at an earlier age than was the case 15 to 20 years ago. Nevertheless, it remains unknown whether Durkin's prediction of a larger population of early readers is accurate.

Method

Purpose of the Study

Answers were sought to the following questions: Prior to any formal school instruction in reading, what percentage of children first entering kindergarten in the Saskatoon Public School System, where English is the language of instruction, can be identified as early reader (ERS); what are the differences, if any, in the family backgrounds of early readers and nonearly readers (NERS); and can we differentiate early readers from nonearly readers in their book orientation skills?

Instruments

In September of 1984, three screening devices were used to identify ERS and NERS. The first device, a *Teacher Questionnaire*, requested kindergarten teachers to identify first-time kindergarten entrants who were reading 20 or more words. A second screening device, the *Durkin Word List* was developed by Durkin (1966) for her study of ERS and consists of 37 words and six sentences commonly found in basal readers. Children who could correctly pronounce 18 or more words within approximately 20 seconds were considered to be potential early readers while those unable to meet these criteria were considered to be NERS. The final screening device was the *Durrell Analysis of Reading Difficulty*, Oral and Silent Paragraphs (1980). Potential early readers were asked to read both the oral and silent paragraphs and answer the comprehension questions. Those children who answered three or more questions correctly for each paragraph were considered to be ERS.

Two measurement instruments, the *Book Orientation Scale* and *Parent Questionnaire* were used to investigate possible differences in book orientation skills and family backgrounds between ERS and NERS. The *Book Orientation Scale* was used in conjunction with the children's book *Your Turn Doctor* (Perez & Robinson, 1982). This scale, based on the *Sand Scale* (Clay, 1972b), and the research of Goodman and Altwerger (1981) was developed by the investigators. It was designed to measure a child's knowledge about the conventions, terminology, and methodology used in describing and reading a book. The *Parent Questionnaire* was also compiled by the researchers and consisted of 45 questions concerning the demographic characteristics of the family, reading habits of the family, home learning experiences including television viewing, travel, preschool attendance, informal instruction in printing, reading and mathematics, and recreational and educational interests of the subject. Each item on the *Parent Questionnaire* gave a number of

possible answers and provided space for additional choices or comments. The *Parent Questionnaire* was piloted in August 1984, prior to commencement of this study.

Procedure

In September 1984, all the 44 kindergarten teachers in English instruction classrooms within the Saskatoon Public School Division (N=1405 students) were asked to complete the *Teacher Questionnaire*. During October 1984, the identified children (13 males and 27 females) were administered the *Durkin Word List* and the Oral and Silent Paragraphs of the *Durrell Analysis of Reading Difficulty* (1980) within their school by the researcher. Twenty-three ERS were identified. A random selection of 20 ERS was made and 20 nonearly readers (NERS) were matched by classroom. The ERS (eight males and 12 females) and the NERS (12 males and eight females) were administered the *Book Orientation Scale* by the researcher within the subject's school. Testing time did not exceed 35 minutes. The caregivers of the randomly selected ERS and NERS were then mailed the *Parent Questionnaire*. *Parent Questionnaires* were completed and returned by the caregivers of 16 ERS and 15 NERS. Interviews were conducted with four ERS' caregivers and four NERS' caregivers who were selected randomly. All data were gathered by December 1984. Data were analyzed using either a t-test, chi-square, or statement of frequency of responses. Extensive profiles were developed for the ERS and NERS interviewed.

Results

Percentage of Population Identified as ERS

Early readers represented 1.64% of the population of students who were enrolled in kindergarten for the first time and who were attending English instruction classrooms. This supports findings by Clark (1976) and King and Friesen (1972) and is somewhat lower than that found by Durkin (1966), Harty (1975), and Hochstetler (1975).

Demographic Differences Between ERS and NERS

Statistically significant differences ($p<0.05$) were found between ERS and NERS on the following demographic features: (a) significantly more ERS had another adult living in the home in addition to their male and female caregivers than did NERS ($t=2.251$); (b) the achieved educational level of the ERS' male caregivers were significantly higher than those of the NERS' male caregivers ($t=2.498$); (c) the ERS' male caregivers were employed at significantly higher occupational levels, as classified by Statistics Canada (1982), than were NERS' caregivers ($t=2.218$). ERS' caregivers were more likely to be employed as professionals or semiprofessionals (62.4%) than were NERS' male caregivers (33%).

Differences in Home Learning Experiences Between ERS and NERS

Early readers were found to differ ($p<0.05$) from NERS in their home learning experiences in the following ways: (a) ERS' male and female caregivers read magazines with significantly greater frequency than did the NERS' caregivers ($t=2.801$); (b) the frequency with which the ERS' female caregiver read the newspaper was significantly greater than that of the NERS' female caregiver ($t=2.802$); (c) ERS

began to copy or write letters at a significantly earlier age than did NERS $\bar{X}$ in months: ERS=32.6; NERS=42.6; $t=2.454$); (d) ERS watched significantly more television per day than did NERS ($\bar{X}$ in minutes per day: ERS=116; NERS=70; $t=2.086$); (e) significant differences were found in the reported benefits of television viewing with more ERS' caregivers stating that television was a medium by which their children learned about the world ($t=2.186$) and increased their vocabularies ($t=2.374$) than did NERS' caregivers; (f) significant differences were found between what the ERS' and NERS' caregivers felt had influenced their children's interest in reading. More ERS' caregivers felt watching television ($t=2.515$), their children's curiosity about the words ($t=3.546$), and telling stories ($t=2.158$) had influenced their child's interest in reading than did NERS' caregivers; (g) the caregivers of ERS gave their children significantly more help in three areas relating to learning to read than did the NERS' caregivers. They discussed letter sounds ($t=2.086$), helped with spellings ($t=2.4998$), and identified words ($t=2.498$) more frequently than did the NERS' caregivers; (h) when caregivers were asked to give three descriptors of their children, significantly more ERS' caregivers termed their children as "shy, introverted, or self-centered" (50% to 13%) and as "curious" (50% to 13%) than did NERS' caregivers. In contrast, significantly more NERS' caregivers termed their children as "loving, caring, kind, or perceptive" (67% to 25%) than did ERS' caregivers.

Book Orientation Skills

ERS were found to be significantly ($p<0.05$) superior to NERS in their book orientation skills. Specifically, ERS were significantly better at locating the beginning of the story within the book ($t=2.690$), showing left-right reading patterning ($t=3.794$) identifying the beginning ($t=6.294$), and ending of the story ($t=6.294$), identifying the authors of the story ($t=2.847$), identifying a word ($t=4.127$), identifying a capital letter ($t=2.333$), recalling detail ($t=3.509$), comprehension of the story ($t=3.509$), and knowledge of the term "last" ($t=2.626$). The data are consistent with previous research which found NER kindergarten entrants to be unfamiliar with the physical act of reading and the terminology commonly used in describing books (Barney, 1976; Clay, 1972a; Downing, 1970; Reid, 1966).

Interviews with Caregivers

Interviews with caregivers of four ERS and four NERS supported the data and also indicated the following differences between ERS and NERS: (a) ERS' caregivers, especially their mothers, appeared to have labeled objects within the environment, written notes, numbers, and words more often for their children than did NERS' caregivers; (b) although the caregivers of ERS did not intentionally set out to teach their children to read, they held and expressed a positive attitude towards their child's potential success with learning to read more frequently than did the NERS' caregivers. This positive attitude appears to have encouraged the development of a stimulating and learning oriented environment within the ERS' home. As well, the caregivers' positive attitude towards their child's future success with reading and learning in general may have caused the caregivers of ERS to set higher expectations for their children than were set by the caregivers of NERS; (c) all ERS appeared to have a great facility for learning and remembering. Learning was viewed as a challenge, but also as a "fun-filled" experience; (d) ERS appeared to be self-motivated and showed great initiative in their desire to learn not

only about reading, but also about mathematics, writing, and the world around them; ERS' caregivers considered them to be curious children; (e) ERS appeared to use television mainly as an information source, and secondarily as a source of reinforcement of skills already learned and as a source of entertainment. This is in contrast to NERS who appeared to use television mainly as a medium of entertainment and secondarily as a means of learning about the world; (f) the ERS appeared to be less interested in active, physical activities than were NERS. Any interest shown by ERS in physical activities was considered to have been stimulated through peers.

Discussion

Durkin (1978) asserted that more and more preschoolers were learning to read at an earlier age than was the case 15 to 20 years before. However, in this study readers represented 1.64% of the total population of kindergarten classrooms where English was the language of instruction. This supports previous findings (Briggs & Elkind, 1977; Clark, 1976; King & Friesen, 1972) and is somewhat lower than found by Harty (1975) and Hochstetler (1975). One might hypothesize that in Saskatoon the percentage of ERS in non-English instruction classrooms may be slightly higher. This hypothesis is based on the fact that most French and Ukrainian Immersion classes in this city are located in upper socioeconomic areas where caregivers are educated at a higher academic level and are presently employed in higher status occupations. In addition, since teacher judgment was used as the initial screening device, some early readers may have been missed. Nevertheless, in this particular sample, there does not appear to be a large increase in the number of ERS as was predicated by Durkin (1978).

Demographic differences found in previous research are, for the most part, supported by this study, while previous literature regarding the home learning experiences of ERS is both supported and contradicted (Table 1). One must conclude that additional factors to be discussed below, but not investigated in this and previous research, play a role in determining whether or not a child will become an ER.

Briggs and Elkind (1977) claimed that interest on the part of parents rather than the child may have been the main motivational determinant. The present research indicted that the caregivers of ERS held and expressed a more positive attitude towards their child's future success in reading than did the NERS counterparts. Although this appears to be a factor, one must still question whether parental attitude and expectations, though a necessary factor, represent the major determinant in the development of reading skills in these children, or whether positive parental attitudes, higher expectations, and a stimulating environment act as supportive secondary factors to "the nature of the child as an organism" (Thomas, Chess, Birch, Hertzig, & Korn, 1963).

The question remains, what differentiated the "nature" of ERS from the "nature" of NERS? On the basis of the limited data available in this and previous studies, one can only speculate about possible answers.

Schema theorists suggest that comprehension and learning are a function of the relationship between information and the currently available skills and knowledge of the learner. Bransford (1979) expands on the nature of this relationship stating the "most important factor underlying our ability to learn is the availability of criteria for evaluating the adequacy of our present level of understanding and

TABLE 1

Finding Compared With Previous Literature

Findings	Supportive Literature	Contradictory Literature
1. 1.64 percent of population are early readers.	Briggs & Elkind (1977); Clarke (1976); King & Friesen (1972)	Durkin (1978); Harty (1975)
2. a. ERS' fathers are employed in high status occupations.		
b. ERS come from upper SES groups.		
c. The male caregivers of ERS were better educated than the NERS' male caregiver.	Briggs & Elkind (1977; 1973); Getman (1979); Harty (1975); Hochstetler (1975); King & Friesen (1972); Miller (1969); Witty & Coomer (1955)	Artley (1980); Durkin (1966); Manning & Manning (1984); Salus (1982); Torrey (1969)
3. There is no significant difference between ERS' and NERS' family size.	Clarke (1976)	Durkin (1966); King & Friesen (1972); Salus (1982).
4. a. The ERS' female caregiver read the newspaper more frequently than did the NERS' female caregiver.		
b. The ERS' caregivers read magazines more frequently than did the NERS' caregivers.	Hochstetler (1975); Durkin (1966)	
5. Significantly more ERS had an adult besides their male and female caregiver living in the home.		
6. ERS watched signficantly more television than NERS.		Clark (1976); Durkin (1966); Harty (1975); King & Friesen (1972); Manning & Manning (1984)
7. Significantly more ERS' caregivers reported that their children learned about the world and increased their vocabularies through television viewing than did NERS' caregivers.		
8. Imaginative play was less enjoyed by ERS than by NERS.	Clark (1976); Durkin (1966); King & Friesen (1972); Nelts (1965)	
9. ERS began to copy letters at a significantly earlier age than NERS.	Showed an interest in spelling, writing, as well as reading. Clark (1976); Durkin (1966); King & Friesen (1972); Lass (1983).	

Table 1 (continued)

Findings	Supportive Literature	Contradictory Literature
10. Significantly more ERS' caregivers reported that		
a. watching television		
b. their child's curiosity about the world		
c. telling stories influenced their child's interest in reading than did NERS' caregivers.	Clark (1976); Durkin (1966); Krippner (1963); Witty & Blumethal (1957)	
11. ERS caregivers gave significantly more help in		
a. discussing letter sounds	Durkin (1966); King & Friesen (1972)	
b. helping with spellings	Durkin (1966)	
c. idenitifying words	Durkin (1966); King & Friesen (1972)	
than did NERS' caregivers		
12. Signficantly more ERS' caregivers termed then as		
a. shy, introverted or self-centered	Clark (1976); Durkin (1966); Torrey (1969)	
b. curious	Clark (1976); Durkin (1966); Krippner (1963); Witty & Blumenthal (1957)	
than did NERS' caregivers.		
13. ERS are significantly superior to NERS in their book orientation skills.	Carlson (1975)	
14. Kindergarten entrants are not familiar with the terminology associated with describing books and physical acts required to read books.	Barney (1976); Clay (1972); Doakes (1982); Downing (1970); Hardy, Stennett & Smith (1974); Meltzer & Herse (1969); Reid (1966)	
15. There were no significant differences in the amount ERS and NERS were read to by their caregivers.	BOS may moderately develop when encouraged by parents without direct instruction. Bissex (1980); Clay (1972)	Early manifestations of BOS may be acquired through reading and rereading books. Chomsky (1972); Holdaway (1979); Smith (1980)

mastery" (pp. 194-195). Initially, ERS appeared to practice and rehearse concept definitions. This was exemplified in their frequent requests that letters, words, and numbers be written and rewritten and in their exhibition of "reading-like" behaviors as described by Doake (1982) and Huey (1908). According to their caregivers, all of the ERS and most of the NERS exhibited "reading-like" behaviors or "mock reading." However, as shown in the significant differences between ERS and NERS on many of the items on the *Book Orientation Scale*, ERS have a superior knowledge of the act of reading and the terminology associated with the reading of print. Bransford (1979) suggested, "the rehearsal of concept definitions, facilitates remembering but does not increase understanding" (p. 233).

The literature about metacognition emphasizes the importance of analyzing the types of experiences that help people learn to utilize appropriate knowledge when needed and those methods which fail to develop such skills. The home environments of ERS shed some additional light in this area. First the ERS' environments appear to encourage active and spontaneous practice in dealing with new situations. For example the mother of Andrew, an ER, stated

> I almost always take five or ten minutes to talk about or explain what was on the [television] show when it's finished. If there's something Andrew or one of the other chidren want to know about we look it up in the encyclopedia or get a book from the library the next time we're there.

A second example was provided by the father of another ER, Alex. While browsing through a magazine, Alex, aged four, read an equation in a "skill testing question" found in an advertisement. Subsequently he asked his father how to complete the computations. Alex's father responded by not only showing Alex the mathematical procedures to be completed, but he also provided a means, in this case a calculator, by which Alex could successfully complete the mathematical computations and solve the equation. At the time of this study, Alex, aged 61 months, was writing and developing his own equations using both an x and y variable, and he was successfully solving these equations using the calculator. Thus it appears that if an ER did not understand a concept presented on television or in print, he or she was encouraged, shown, modeled, and provided with a practical means for solving gaps in knowledge.

Secondly, ERS live in an environment where feedback is very much the norm. At a metacognitive level, ERS unlike NERS appear to be able to monitor accurately their current levels of understanding and mastery. They also appear to be able to take action through questioning their parents and seeking information from television, or from print, to raise their current level of understanding. Their environment appears to encourage action. They are not fearful of being thought foolish for seeking answers to gaps in their knowledge, whether it is in the recognition of a letter, a sound, or a word. Thus the environment of ERS allows for risk-taking with the generating and testing of hypotheses actively encouraged, if not demanded. Levine (1972) suggests individuals who learn best do not appear to passively absorb information; rather they attempt to analyze and process information to form concepts. This is certainly true of ERS who, as was indicated by the interviews, took an active role in their learning, often selecting topics for which they requested additional information. All the ERS asked that books be read and reread to them. Many requested that letters be written and identified. The words of one ER's mother illustrate the strong desire the ERS had to learn.

> Previously, especially between one and four years, he was asking so many questions and he asked me to read all his books . . . he'd ask me to write all the numbers to 100 at least

> four times . . . he'd ask me to read every day first thing in the morning. . . . Initially, he wouldn't even allow me to take a bath; he wanted me to read all his books at once. . . . Even when we went to a friend's house he'd ask me to bring some paper and something to write with so I could write them [the numbers from 0 to 100 and letters].

One might also hypothesize that ERS are more able than NERS to anticipate situations in which their knowledge must be utilized. The profiles of ERS indicated that an ER might use television listings to determine what programs he or she could view, or read the cover of a record album to select the music to which he or she might listen. The television, signs, magazines, and record albums read by ERS became more than just a viewing or listening activity. Rather, ERS used these materials as sources of information and tools for learning more about letters and words. But more importantly, they anticipated potential uses of print, made self-evaluations of their strengths and weaknesses, and determined either that they were adequately prepared or that they required additional information and practice. ERS then set realistic goals for themselves and proceeded to manipulate their environments so that they could reach these goals. For example, if an ER was unfamiliar with several letters, he or she might demand that caregivers identify and write these letters. Their caregivers would then indirectly place them in a number of "test situations" by playing rhyming games, alphabet bingo, or cloze games. Similar strategies were employed with concepts relating to mathematics, words, and the world around them. Furthermore, the caregivers of ERS often raised questions when reading and speaking with their children that stimulated and further expanded their children's knowledge base. The test situations were always low risk, low stress, as well as game activities. Hence learning remained a fun-filled, enjoyable activity.

Implications and Recommendations

Within a structured learning environment, teachers can apply some of the strategies employed by the caregivers of ERS. As Durkin (1978) and Clay (1980) have suggested, teachers should read to and listen to children. Furthermore, adults should be reminded to speak with, as opposed to at, children. Teachers can and should model reading, not only as a leisure activity but also as a means by which questions about the world can be answered. Strategies and steps for seeking information need to be discussed as they are modeled. Rather than suggesting additional information may be found in a reference book, teachers should model and discuss not only where this information may be found, but also how new information relates to what the child already knows.

Just as the ERS' homes were child-centered, so too should be the classroom learning environment. There should be opportunities for self-initiated and self-directed learning from highly interesting materials.

The homes of ERS allowed for and encouraged practice activities that were highly motivational, game-like, manipulative, and social in nature. Inherent in the practice activities of the ERS was the assumption that if learning is occurring, errors and mistakes are acceptable, if not necessary. Risk taking and nonthreatening "testing" of newly acquired learning was, at the very least, encouraged and at the most, demanded by the caregivers of ERS. Expectations on the part of the caregivers were always positive in nature; ERS were viewed as children who could and would learn. As McCracken and McCracken (1985) suggest, the use of practice activities where errors are acceptable, can and should be implemented into classrooms.

Suggestions for Further Research

Major gaps in our knowledge base exist with respect to the everyday social interaction of ERS and their caregivers. Since the families of several ERS contained multiple ERS, the younger siblings of identified ERS provide a population which may shed additional light in this area.

Secondly, ERS have been termed "supercurious," "self-motivated," and "self-directed" learners. The caregivers in this study restate that they, like parents in previous studies of ERS, began direct and indirect instruction at the request of their children. An explanation ought to be sought about what initiated this intrinsic motivation or curiosity in ERS, and why it is either not found or not initiated in all children.

Conclusion

In summary, the environment of ERS appears to have been conducive to learning. Caregivers of ERS described their children as truly active learners, having a sense of what they needed to do in order to facilitate their own understanding of reading and learning in general. These caregivers provided an environment that was conducive to and highly supportive of asking questions for clarification; an environment with caregivers who modeled and reinforced the development of strategies for completing gaps in knowledge; and an environment in which fearless practice and self-monitoring of comprehension was facilitated. The environment created by these caregivers encourages concept development, skill development, and the development of criteria for guiding learning. According to the caregivers of ERS, they taught their children not how to read, but how to learn. The question of who or what initially triggered the ER's first interactions with his or her environment, and whether this interaction was due to the nature of the child or to the nature of the environment, remains unanswered.

References

Barney, D. (1976). Kindergarten four-year-olds' concept of reading, knowledge of books and print. In D.B. Doake & B.T. O'Rourke (Eds.), *New directions for reading teachers*. Wellington: New Zealand Educational Institute.

Bransford, J.D. (1979). *Human cognition: Learning, understanding, remembering*. Belmont, CA: Wadsworth.

Briggs, C., & Elkind, D. (1973). Cognitive development in early readers. *Development Psychology, 9*, 279-280.

Briggs, C., & Elkind, D. (1977). Characteristics of early readers. *Perceptual and Motor Skills, 44*, 1231-1237.

Brown, C.R. (1985). *Precocious readers: A comparative study of early reading and non-early reading kindergarten entrants*. Unpublished master's thesis, University of Saskatchewan.

Carlson, J.R. (1975). *A study of selected reading perceptions of kindergarten children who were early readers*. Unpublished doctoral dissertation, University of Nebraska.

Clark, M. (1976). *Young fluent readers*. London: Heinemann.

Clay, M. (1980). *The early detection of reading difficulties: A diagnostic survey* (2nd ed.). New York: Heinemann.

Clay, M.M. (1972a). *Reading: The patterning of complex behavior.* Auckland: Heinemann.

Clay, M.M. (1972b). *Sands: Concepts about print test.* Auckland: Heinemann.

Doake, D. (1982). *Book experience and emerging reading behavior in preschool children.* Unpublished doctoral dissertation, University of Alberta.

Downing, J. (1970). Children's concepts of language and learning to read. *Educational Research, 12,* 106-112.

Durkin, D. (1966). *Children who read early.* New York: Teachers' College Press.

Durkin, D. (1978). *Teaching children to read* (2nd ed.). Boston: Allyn and Bacon.

Evans, J.R., & Smith, L.J. (1976). The psycholinguistic skills of early readers. *Reading Teacher, 30,* 39-43.

Getman, M. (1979). *Correlates of early reading.* Unpublished doctoral dissertation, University of Minnesota.

Goodman, Y.M., & Altwerger, B. (1981, September). *Print awareness in pre-school children: A working paper. A Study of the Development of Literacy in Preschool Children.* Arizona Center for Research and Development, University of Arizona.

Harty, K.F. (1975). *A comparative analysis of children who enter kindergarten reading and of children at the same age who require additional readiness for reading.* Unpublished doctoral dissertation, University of Wisconsin.

Hochstetler, M.E. (1975). *A study of factors related to the reading abilities of beginning kindergarten children.* Unpublished doctoral dissertation, Ball State University.

Huey, E.B. (1908). *The psychology and pedagogy of reading.* New York: Macmillan.

King, E.M., & Friesen, D.T. (1972). Children who read in kindergarten. *The Alberta Journal of Educational Research, 18,* 147-161.

Levine, M. (1972). *Hypothesis testing, a cognitive theory of learning.* Hillsdale, NJ: Erlbaum.

Manning, M., & Manning, G. (1984). Early readers and non early readers from low socioeconomic environments: What their parents report. *Reading Teacher, 38,* 32-35.

McCracken, R.A., & McCracken, M.J. (1985). *Teaching children to read.* Paper presented at Saskatchewan Reading Conference, Saskatoon.

Perez, C., & Robinson, D. (1982). *Your turn doctor.* New York: Dutton.

Plessas, G.P., & Oakes, C.R. (1964). Prereading experiences of selected early readers. *Reading Teacher, 17,* 241-245.

Reid, J.F. (1966). Learning to think about reading. *Educational Research, 9,* 56-62.

Thomas, A., Chess, S., Birch, H.G., Hertzig, M.E., & Korn, S. (1963). *Behavioral individuality in early childhood.* New York: New York University Press.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 234-248.*

ROBERT B. KOTTKAMP
Hofstra University

and

ARTHUR L. TRAVLOS
Scotch Plains-Fanwood School District, NJ

Selected Job Stressors, Emotional Exhaustion, Job Satisfaction, and Thrust Behavior of the High School Principal[1]

This study investigated the relationships between four perceived role related stressors (role conflict, role ambiguity, role overload, and powerlessness), two affective reactions to work (emotional exhaustion and job satisfaction), and thrust leadership behavior among high school principals. Principals from 74 New Jersey high schools completed surveys containing measures of role stressors and responses to work. A random sample of teachers in each school reported on the principal's thrust behavior. All four stressors were significantly correlated with the two affective responses to work. However, the stressors explained more variance in emotional exhaustion (R^2=.65) than in job satisfaction (R^2=.39). Neither the stressors nor the affective responses were related to principal thrust except for the age range of 40-46.

Over the previous decades, educators have been beset with a multitude of problems. However, no problem may prove more debilitating than one confronting them now: job dissatisfaction, emotional and physical exhaustion, and the many other psychological, physical, and behavioral manifestations related to job stress or "burnout." Yet while stress and burnout have become established topics in educational research in recent years, teachers have received more attention than administrators. In a recent and extensive review of Canadian research on teacher stress, for example, Hiebert (1985) included a short "Note on School Administrators," which referenced only three empirical studies. The purpose of this study was to add to the small stock of conceptually based empirical research on stress among school administrators; the focal role was that of the high school principal.

Dr. Kottkamp is Associate Professor of Educational Administration whose research interests include occupational stress, organizational analysis, the principalship, and the sociology of teaching.

Dr. Travlos was recently appointed Director of Staff Development and is presently conducting an instructional skills program for teachers.

Not only are there few studies of stress among principals, but the results of existing studies conflict. Especially problematic is the level of stress reported among principals. On the one hand, there are reports of considerable stress. Lusthnus (1975), for example, found that principals in Quebec experienced more pressure in their jobs than teachers or other professional groups. Gmelch, Koch, Swent, and Tung (1982) conducted a survey of all Oregon school administrators from vice principal through superintendent. "More than 60% reported that at least 70% of their total life stress resulted from their jobs" (p. 11). Swent and Gmelch (1977) found administrators reported 75% of their total life stress came from their jobs. Syrotuik and D'Arcy (1983), in a study of administrators from Saskatoon, reported that role conflict was the most important source of stress. On the other hand, MacPherson's (1985) study of principals in Nova Scotia showed low levels of role conflict, role ambiguity, and emotional exhaustion and mild to moderate levels of overall job stress. In an Alberta study, Jankovic (1983) reported only mild to moderate levels of stress among 88% of principals sampled. These conflicting results among the small number of existing studies are probably related to the differing conceptual frameworks and operational definitions used in the various inquiries.

Given the mixed results to date, several reasons remain for pursuing studies of principal job-related stress and its effects. First, there is a concern for the individual occupying the role. It is important, both theoretically and practically, to understand the types and amounts of stress inherent in the school leadership role and what the effects of these are likely to be for the individual. The rationale for this concern is self-evident, and all the empirical studies of administrator stress referenced in this study address aspects of this issue. Second, because of the principal's hierarchical position, stress is likely to have an organizational dimension as well as a strictly individual one. Caldwell and Forney (1982), for example, note that "the principal has been considered as the most influential factor in determining the climate, direction, and leadership of a school" (p. 1). Erez and Goldstein (1981), in a study of Israeli principals, argued that experienced stress is a major factor affecting principal leadership efficiency. Thus, because of the principal's central position in the school's social system, debilitating effects of stress on the individual may be translated into larger negative effects for the whole school.

Problem

This study investigated the relationships between perceived role related job stressors, internal affective responses to work, and a leadership behavior among high school principals. Three organizational factors (role conflict, role ambiguity, and role overload) and a dimension of alienation (powerlessness) were examined to determine their relationships with two affective responses to work: job satisfaction and emotional exhaustion, a component of Maslach and Jackson's (1978) burnout syndrome. Moreover, these four job stressors and two affective conditions were analyzed to determine their relationships with a leadership behavior presumed to have organizational consequences, the thrust behavior of the principal. Thrust is the principal's attempt to motivate teachers through setting a personal example.

Theoretical Framework

Of the four job related stressors examined in this study, three have their conceptual roots in role theory. Role conflict refers to a situation in which an individual perceives incompatible work demands. Kahn, Wolfe, Quinn, and Snoek (1964) reported that role conflict was related to various consequences of job stress including job dissatisfaction, high job-related tension, greater sense of futility, and lower self-esteem.

Role ambiguity exists when a person perceives inadequate information about expectations for his or her work. It has been linked to job dissatisfaction, poor production, and anxiety (Kahn, Wolfe, Quinn, & Snoek, 1964), increased blood pressure and pulse rate (French & Caplan, 1973), depression and resentment (Caplan & Jones, 1975), lower self-esteem and life dissatisfaction (Margolis, Kroes, & Quinn, 1974).

Role overload refers to a perception of more work than a person can complete in a given time. Kraut (1965) and Mueller (1965) reported a variety of negative reactions in individuals exposed to overload: tension, low job satisfaction, poor interpersonal skills, and low self-esteem. Moreover, French and Caplan (1973) found it to be related to absenteeism, low motivation to work, and an absence of suggestions to employers.

The fourth precursor of job stress examined in this study is one of the five dimensions of alienation conceptualized by Seeman (1959). Powerlessness is the perception that an individual cannot control the outcomes he or she seeks. It has been linked to job dissatisfaction (Shepard, 1972), poor job performance and tardiness (Cummings & Manring, 1977), lower self-esteem and increased anxiety (Gilbert & Mangelsdorff, 1979).

Emotional exhaustion and job satisfaction were the two affective responses to work examined in this study. Emotional exhaustion is the feeling that one's emotional resources are depleted; that one is no longer able to give to others on the psychological level. Associated with emotional exhaustion is a sense of physical exhaustion and a lack of any positive feeling, sympathy, or respect a professional would normally show for clients. It is a subset of an overall burnout index developed by Maslach and Jackson (1978). Researchers have found positive relationships between emotional exhaustion and the following: low morale, high absenteeism, and turnover (Maslach & Jackson), cynicism, negativism, inflexibility, and a sense of helplessness and hopelessness about one's clients (Freudenberger, 1977).

Job satisfaction is the affective evaluation an individual experiences toward a job. It has been negatively related to the following outcomes: absenteeism (Porter & Steers, 1973), turnover (Waters & Roach, 1971), early retirement (Vroom, 1964), tardiness and grievances in organizations (Mirvis & Lawler, 1977), and poor mental health (Kornhauser, 1965). Friesen, Holdaway, and Rice (1984) noted that most studies of job satisfaction have focused on production workers rather than administrators. They argued the need for more studies of administrators and specifically suggested investigating the relationship between job stress and dissatisfaction.

The general conceptual link between role based stressors in the work place and negative affective responses to work was established and first empirically tested by Kahn et al. (1964). More recently these relationships have been confirmed in studies of several groups of educators. Role conflict and ambiguity have been correlated with perceived emotional exhaustion among several general teacher samples (Schwab & Iwanicki, 1982; Beck & Gargiulo, 1983), special education teachers

(Beck & Garguilo, 1983), and high school department chairpersons (Kottkamp & Mansfield, 1985). Role conflict, but not role ambiguity, has been related to emotional exhaustion among principals (MacPherson, 1985). Perceived powerlessness was also correlated with emotional exhaustion among secondary department chairpersons (Kottkamp & Mansfield, 1985). Role conflict and ambiguity were correlated with job dissatisfaction and job-related tension among multidisciplinary team administrators (Kottkamp & Helfman, 1985). Role overload, a separate concept but related to role conflict and ambiguity, has yet to be empirically tested as a stressor among educators. However, it is reasonable to assume that it is related to affective responses to work in ways which parallel those of the other three stressors for which evidence already exists.

Taken together, this conceptual and empirical background leads to two hypotheses:

H_1 There are negative relationships between the perceived role conflict, role ambiguity, role overload, and powerlessness of the high school principal and his or her job satisfaction.

H_2 There are positive relationships between the perceived role conflict, role ambiguity, role overload, and powerlessness of the high school principal and his or her emotional exhaustion.

The first two hypotheses were designed to test relationships between role stressors and two *affective* responses: job satisfaction, and emotional exhaustion, at the individual level. However, there was also interest in pursuing an independently measured *behavioral* outcome of the stressors and the two affective responses, a behavioral outcome which would be likely to effect the *organization* as well as the individual.

The literature indicated that the job stressors and affective responses in this study have been related to various behavioral outcomes. Role conflict and ambiguity have been associated with decreased productivity (Kahn et al., 1964). Role overload has been correlated with poor interpersonal relations (Kraut, 1965; Mueller, 1965), and high absenteeism (French & Caplan, 1973). Powerlessness has been related to poor performance and increased tardiness (Cummings & Manring, 1977). In a similar vein, negative affective responses to work have been related to the following negative behaviors. Job dissatisfaction has been correlated with increased absenteeism and tardiness (Mirvis & Lawler, 1977; Waters & Roach, 1971), while emotional exhaustion has also been associated with higher absenteeism, job turnover (Maslach & Jackson, 1978), and cynicism and negativism toward clients (Freudenberger, 1977).

A search was carried out to identify a principal behavior with organizational consequences, a behavior which could also be related to the stressors and affective responses of job satisfaction and emotional exhaustion. The leadership behavior, thrust, was chosen. Halpin and Croft (1962) defined thrust as a dynamic behavior of the principal characterized by efforts to "move the organization." Though "starkly task-oriented," this behavior is not marked by close supervision. Rather, the principal asks no more of teachers than he or she is willing to give and leads through active personal example and modeling of desired behavior.

High thrust behavior is an essential component of open school climate as conceptualized by Halpin and Croft (1962). Climate openness has, in turn, been associated with a number of other variables typically considered to be desirable: higher teacher satisfaction (Schwandt, 1978) and loyalty to the principal (Kanner,

1974); greater teacher confidence in their own and the school's effectiveness (Andrews, 1965); higher humanistic Pupil Control Ideology scores (Appleberry & Hoy, 1969); greater innovativeness (Thomas, 1976); and lower student alienation, higher student morale, and lower dropout rates (Hartley & Hoy, 1972; Mullins, 1976). Because scholars have argued that the principal is the key figure in setting the school climate (Drummond, 1970; Fox, 1973; Hoy & Henderson, 1983) and because the correlates of open climate are typically considered positive, the principal with high thrust behavior is likely to have a positive effect on his or her school while the principal with low thrust behavior is likely to have a negative effect.

Most of the individual outcomes of role stressors and their negative behavioral responses such as decreased productivity, absenteeism, tardiness, job turnover, and negativism towards clients may be categorized as withdrawal behaviors. These behaviors indicate either a physical or a psychological withdrawal from central elements of effective work performance. Quite the opposite of withdrawal from work is thrust, a leadership behavior indicated by such activities as high energy expenditure and modeling, a leadership behavior which has beneficial consequences at the organizational level. Thus, if role stressors and their negative affective responses have been associated with various withdrawal behaviors in other work settings, it is logical to predict that stressors and their negative responses will be related to reduced energies put into the work of the principalship. This reasoning led to the hypothesis:

H_3 Perceived role conflict, role ambiguity, role overload, powerlessness, job dissatisfaction, and emotional exhausion are negatively related to the principal's thrust behavior.

Method

Sample

Data for the study were collected by means of questionnaires distributed to the principals and teachers in 74 public high schools in New Jersey. Although participation was voluntary, an effort was made to obtain a reasonably representative sample of high schools in the state. Sample schools had student populations ranging from 250 to 2000, were drawn from 17 of 22 New Jersey counties, and were drawn from urban, suburban, and rural areas. The sample schools also covered the whole range of 10 District Factor Groups (DFGs) established by the New Jersey State Department of Education. DFGs are computed as a weighted composite of district averages of seven socioeconomic and demographic factors: educational and occupational background, per capita income, poverty and unemployment rate, population density and mobility, and degree of urbanization. The sample schools were skewed slightly toward higher DFG ratings or higher socioeconomic status when compared to the state average.

The principals completed scales to measure the four job stressors and two affective responses to work. The data collected for this study were part of a set of related studies of high schools which in total required four independent teacher questionnaires. Therefore, a randomly selected one-fourth of the teaching faculty of each school was administered a scale to measure the thrust behavior of the principal. Because teacher surveys were completed in faculty meetings, usable return rates exceeded 90%.

The unit of analysis was the individual high school principal in the context of his or her particular school and district. Teacher responses on the thrust scale were

aggregated to form a school level perceptual score. The data were analyzed through correlation and regression analyses. Secondary analyses included t-tests and Fisher Z transformations.

Instrumentation

Role Conflict, Ambiguity, and Overload. Role conflict, ambiguity, and overload were measured by a 15-item instrument (5 items for each role stressor) refined by Abdel-Halim (1978) from the previous work of Beehr (1976) and Rizzo, House, and Lirtzman (1970). In Abdel-Halim's instrument, each respondent is asked to indicate on a 7-point scale ranging from "very false" to "very true," the extent to which the items described his or her work situation. For these scales and the measures of powerlessness, job satisfaction, and emotional exhaustion, a consistent set of 6-point response scales with poles of "strongly agree" and "strongly disagree" were used. This modification was made so that all six measures could be intermixed in the survey to reduce response bias.

Construct validity of the instrument was established by Beehr (1976) and Rizzo, House, and Lirtzman (1970). Reliability coefficients of .76 (conflict), .69 (ambiguity), and .74 (overload) were reported by Abdel-Halim (1978) using the Spearman-Brown formula. In this study the alpha coefficients were .70 (conflict), .62 (ambiguity), and .81 (overload).

Powerlessness. Powerlessness was measured by a 12-item subscale of an instrument developed by Zeilinski and Hoy (1983). The subscale measures perceived powerlessness in the school as a whole or what the researchers called Organizational Effect. Construct validity for the instrument was assessed by Zeilinski and Hoy. High reliability was reported for the 12 items, an alpha coefficient of .92. The data in this study generated an alpha coefficient of .88.

Emotional Exhaustion. This concept was measured through one subscale of the Maslach Burnout Inventory (MBI) developed by Maslach and Jackson (1978). The MBI is a measure of perceived burnout along three dimensions: emotional exhaustion, depersonalization, and personal accomplishment. The original instrument consists of 22 items which are rated twice, once for frequency and once for intensity. For this study, only the 9-item emotional exhaustion subscale was used. Iwanicki and Schwab (1981) had suggested collapsing the frequency and intensity responses into a single response category because of the high correlations (.89 for emotional exhaustion) between the two when used in a study of teachers. Consequently, a single 6-point "strongly agree" to "strongly disagree" scale was used in this study.

Maslach and Jackson (1979) reported alpha reliability coefficients for the emotional exhaustion subscale of .89 (frequency) and .87 (intensity). In this study the alpha coefficient for the modified emotional exhaustion scale was .88. Several investigators (Iwanicki & Schwab, 1981; Maslach & Jackson, 1979) have demonstrated the validity of the MBI through concurrent, discriminant, and construct methods.

Job Satisfaction. The concept of job satisfaction was measured by the 6-item Index of Satisfaction developed by Miskel, Fevurly, and Stewart (1979). They provided evidence of content validity for the instrument and reported an alpha coefficient of .78. The alpha in this study was .71.

Thrust Behavior. This principal behavior was measured using a scale similar to the thrust subscale of the Organizational Climate Description Questionnaire

(OCDQ) developed by Halpin and Croft (1962). As part of the work to revise the OCDQ for secondary school use (Kottkamp, Mulhern, & Hoy, 1985) the original thrust subscale was modified by adding two items to the original 9-item scale. The new 11 item thrust subscale asks respondents to reply to statements about the principal's behavior on a 4-point scale ranging from "rarely occurs" to "very frequently occurs." The alpha coefficient for the revised thrust subscale reached .93 in this study.

Results

Descriptive Statistics

The scale and item means, standard deviations, and zero order correlations for the variables in the study are presented in Table 1. All principal self-report variables were measured on 6-point scales. The conceptual mean for a 6-point scale is 3.5. As may be seen in Table 1, mean item responses of principals to job satisfaction (4.56) were comparatively high, while both role conflict (3.46) and role overload (3.49) hovered right at the conceptual mean. Emotional exhaustion (2.26) was comparatively low at more than one point below the conceptual mean, and powerlessness (1.88) showed the lowest mean item response of all the variables. Principal thrust was reported by teachers on a 4-point scale with a conceptual mean of 2.5. The mean item response (1.45) indicated that teachers perceived relatively low thrust from their principals.

TABLE 1

Means, Standard Deviations, and Intercorrelations

	Variable	1	2	3	4	5	6	Item Mean	Scale Mean	Scale *SD*
1.	Role Conflict							3.46	17.30	5.07
2.	Role Ambiguity	.36*						2.24	11.32	3.31
3.	Role Overload	.46*	.40*					3.49	17.47	5.55
4.	Powerlessness	.44*	.50*	.57*				1.88	22.55	8.26
5.	Job Satisfaction	−.47*	−.44*	−.44*	−.59*			4.56	27.34	5.13
6.	Emotional Exhaustion	.57*	.40*	.68*	.72*	−.61*		2.26	20.38	8.22
7.	Thrust Behavior	−.14	−.14	−.12	−.05	.01	−.07	1.45	15.98	2.28

*$p<.01$

Several patterns are evident among the correlations. First, the three role perception variables and perceived powerlessness are all related at moderate levels, with correlations ranging from .36 (conflict with ambiguity) to .57 (overload with powerlessness). Thus job stressors in this study share proportions of variance ranging from one-eighth to one-third. These proportions of shared variance indicate both relatedness and relative independence, as has been found in previous studies (Abdel-Halim, 1978; Rizzo, House, & Lirtzman, 1970). Clearly, the job stressors are not tautological, but their relatedness is likely to result in some collinearity when the affective and behavioral variables are regressed on them.

The second pattern is that all job stressors are negatively correlated with job satisfaction and positively correlated with emotional exhaustion at statistically significant levels. These relations range from .40 (ambiguity with emotional exhaustion) to .72 (powerlessness with emotional exhaustion). These relationships suggest support for Hypotheses One and Two. However, a third pattern is that none of the job stressors, job satisfaction, or emotional exhaustion have significant relationships with thrust behavior. These correlations suggest the rejection of Hypothesis Three.

Hypothesis Testing

To gain a more thorough understanding of the individual and combined effects of the variables in the hypothesized relationships, simultaneous inclusion, multiple linear regression analyses were conducted. Table 2 reports the results of regressing satisfaction on the four job stressors, a test of Hypothesis One. The total equation reaches significance, indicating that taken together the four independent variables do predict job dissatisfaction. Further, the adjusted R^2 indicates that slightly over one-third of the variance in satisfaction is related to the variance in the four job stressors. Scrutiny of the betas and their significance level indicates, however, that only role conflict and powerlessness are individually significant predictors of job satisfaction when all four independent variables are simultaneously related to the dependent variable. Hypothesis One receives support, but the regression analysis clarifies the relative strength of the relationships between the job stressors and satisfaction when the stressors are considered jointly.

TABLE 2

Summary for the Regression of Perceived Principal Job Satisfaction on Four Perceived Job Stressors

Independent Variables	*b*	beta	Standard error beta	*t*	Adjusted R^2
Role Conflict	−.23	−.22	.10	−2.06*	
Role Ambiguity	−.22	−.14	.16	−1.31	
Role Overload	−.06	−.06	.11	−.55	
Powerlessness	−.24	−.38	.07	−3.14**	
	Total Equation			F=12.42**	.385

*$p<.05$
**$p<.01$

The results of regressing emotional exhaustion on the four job stressors are presented in Table 3. Again, the total equation reaches significance, but the adjusted R^2 of .65 indicates that the four independent variables explain two-thirds of the variance in emotional exhaustion. The four job stressors, thus, explain twice as much variance in emotional exhaustion as in job satisfaction. The betas indicate that role conflict, overload, and powerlessness are individually significant predictors of exhaustion. Hypothesis Two also receives solid support.

TABLE 3

Summary for the Regression of Perceived Principal Emotional Exhaustion on Four Perceived Job Stressors

Independent Variables	*b*	beta	Standard error beta	*t*	Adjusted R^2
Role Conflict	.36	.22	.13	2.74**	
Role Ambiguity	−.09	−.04	.20	−.43	
Role Overload	.49	.33	.13	3.72**	
Powerlessness	.45	.46	.09	4.95**	
	Total Equation			F=34.42**	.647

**$p<.01$

As could be predicted from the correlations in Table 1, the regression of thrust behavior on the four job stressors, job satisfaction, and emotional satisfaction did not prove significant. The results of this analysis are shown in Table 4. As may be seen, neither the equation nor any individual independent variable explains a significant amount of variance in thrust. Hypothesis Three is soundly rejected.

TABLE 4

Summary for the Regression of Principal Thrust Behavior on Four Perceived Job Stressors and Two Affective Responses

Independent Variables	*b*	beta	Standard error beta	*t*	Adjusted R^2
Role Conflict	−.06	−.13	.07	−.89	
Role Ambiguity	−.09	−.14	.10	−.97	
Role Overload	−.04	−.09	.07	−.56	
Powerlessness	.01	.05	.05	.26	
Job Satisfaction	−.05	−.12	.07	−.75	
Emotional Exhaustion	.00	.02	.06	.02	
	Total Equation			F=.522	.041

Secondary Analyses

Secondary analyses were conducted in order to investigate the failure of Hypothesis Three. These analyses were guided by two issues. First, age has been related to differences in perceived stress, but these findings are complex with some studies reporting low stress with advancing age (Indik, Seashore, & Slesinger, 1964; McGrath, 1970), while in boundary spanning positions (Tung & Koch, 1980) and among higher level executives, perceived stress increases with age (Pincherle, 1972). Second, the adult developmental literature indicates that significant stress related changes, often referred to as mid-life crisis, occur in the first half of the fourth decade for males (Gould, 1972; Levinson, 1974; Lidz, 1980).

When age was correlated directly with the seven variables in Hypothesis Three, all correlations were close to zero. Apparently, there are no linear relationships between age and the variables of interest in this study. The next analysis was based on the adult development literature. Principals were placed in two age related groups: 40-46 years of age (N=27) and 47 years of age and above (N=40). Those 39 or under (N=6) were dropped from the analysis because the group was too small to allow for meaningful analysis. T-tests for mean differences between the two age groups were computed across all seven variables in Hypothesis Three. Neither the role stressors, affective responses, nor thrust behavior showed significant mean differences across these age groups.

Finally, age was analyzed as a "moderator variable" (Zedeck, 1971; Arnold, 1982), intervening in the relationships specified in Hypothesis Three. Correlations were computed between variables of interest within each age group and then compared across age groups for significant differences using Fisher's Z transformation. The question answered was: Does the *degree* of the correlation (e.g., between thrust and role conflict) differ significantly between the two age groups? The correlations and the results of the Fisher Z transformations are shown in Table 5. The degree of the relationship across the two age groups was found significant in five of the six relationships, satisfaction-thrust being the only exception. Apparently, although age was not directly related to perceptions of the various stressors or outcomes themselves, it did, in the constructed age categories, moderate the *degree* of relationships between the stressors, emotional exhaustion, and the behavioral outcome, thrust. Scatter plots for each age group were examined to see whether outliers were responsible for spurious correlations. This did not appear to be the case.

TABLE 5

Moderating Effects of Age Groupings on the Correlations Between Role Stressors, Affective Responses, and Thrust Behavior of Principals

Variables Correlated	Principals 40-46 Years (n=27)	Principals 47 Years and Over (n=40)	
	Correlation Coefficients		Z Score
RC—THRUST[a]	−.54	−.03	2.20*
RA—THRUST	−.51	.08	2.44*
RO—THRUST	−.44	.16	2.41*
PO—THRUST	−.49	.18	2.75**
SA—THRUST	.31	−.14	1.77
EE—THRUST	−.45	.15	2.45*

[a]RC=Role Conflict, RA=Role Ambiguity, RO=Role Overload, PO=Powerlessness, SA=Satisfaction, EE=Emotional Exhaustion

*$p<.05$

**$p<.01$

Discussion

The results of this study are in line with those of MacPherson (1985) and Jankovic (1983) in demonstrating that principals are not experiencing high levels of stress, in this case emotional exhaustion. In fact, although comparison is crude because of the modified version of the Maslach Burnout Inventory used in this study, the high school principals showed a slightly lower mean level of emotional exhaustion than any of the four separate samples of teachers compared by Kottkamp and Mansfield (1985). Further, the sample principals responded with a moderately high mean level of job satisfaction, an affective work response possessing a negative correlation (−.61) with emotional exhaustion. These are encouraging findings to report to principals and to those who aspire to the position.

The findings for Hypotheses One and Two, as predicted, replicate a generally consistent pattern of role perception-affective response relationships found in both educational and noneducational settings. Each role related perceptual variable (role conflict, ambiguity, overload, and powerlessness) showed a significant zero order correlation of moderate strength with each affective work response (job satisfaction and emotional exhaustion). However, because the correlations among the role perception variables were roughly of the same order as the correlations between these variables and the affective responses, multicollinearity was a fact that affected the results of the regression equations. The total equations were significant, but not all role variables showed individual significance in the equations when all four role perceptions were considered simultaneously.

MacPherson's (1985) study of elementary and secondary principals in Nova Scotia makes a useful comparison for discussing patterns of relationships between the role variables and emotional exhaustion. In the present study, role overload, powerlessness, and role conflict, in descending order of explanatory power, all showed individually significant relationships with emotional exhaustion in the regression equation. Among MacPherson's principals, role conflict was also significantly related to emotional exhaustion, while role ambiguity was not. Perceived degree of autonomy described as "freedom to operate their schools as they saw fit" (p. 3) was also negatively related to exhaustion. Autonomy, so defined, seems a polar opposite of powerlessness; thus the resulting patterns in both principal studies appear quite similar along two dimensions. MacPherson's study contained no variable comparable to role overload.

The principal occupies a boundary spanning position. Occupants of such positions are susceptible to numerous conflicting demands from both internal and external sources. These conflicting demands were related to principal emotional exhaustion. To be an effective leader of a school and its social system, the principal also needs to hold clear expectations for the role. Principals seemed to find role ambiguity less of a problem in their leadership role and less stress evoking than conflict. The dual position of leader and boundary spanner generates much work. Most emotionally exhausting to principals in this study was the volume of work generated by the role. As a source of exhaustion, role overload was followed closely by a sense of powerlessness to control events in the school, though the mean level of this sense was quite low among the principals in the sample.

Remembering that the level of emotional exhaustion experienced by principals was not high, the four role related variables did explain a high proportion—almost two-thirds—of the variance in emotional exhaustion. With this level of prediction, the results of this study have practical as well as theoretical implications. Role

stressors were not only theoretically related to emotional exhaustion. By contrast, the same four variables explained only half as much variance in job satisfaction. This proportion of variance is still noteworthy. These findings, along with the correlation of −.61 between satisfaction and emotional exhaustion, also provide a first response to Friesen, Holdaway, and Rice's (1984) call for an examination of the relationship between perceived dissatisfaction and stress.

Reducing the negative aspects of role is a goal which would likely bring actual and desirable personal outcomes for high school principals. Such reduction would be more likely to reduce pain than to bring joy. It is, however, unrealistic to suggest that role conflict, ambiguity, overload, and powerlessness can be drastically reduced in the principalship. They come, to some degree, with the proverbial territory.

The failure of Hypothesis Three remains an enigma. Puzzling indeed is the finding that neither role perceptions nor affective responses were generally related to the thrust variable and that age ranges moderated the relationships such that only 40-46-year-old principals demonstrated the predicted relationships. There are only questions, not answers. For those interested in the age group findings, further discussion is available elsewhere (Travlos, 1984).

Several other issues may be related to the failure of Hypothesis Three. First, the operation for thrust may not be a sensitive measure of principal behavior. Not only was the mean for thrust quite low, but the standard deviation indicated little variation among schools. Second, survey research bears the liability of potential response bias when variables are not measured independently. It is possible that the intercorrelations of the principal self-report variables were inflated because of method bias. If this were the case, then the correlations with thrust were lower, since it was the only independently measured variable.

A note on the measurement of role ambiguity may also be helpful to future researchers. The alpha coefficient for ambiguity reached only .62 in this study, indicating a good deal of error variance in the measurement and attenuation of the correlations beween this and other variables. A decision was made to use the shortest scales with reasonable reliabilities in order to reduce the length of the survey—since surveys do contribute to role overload for principals! Alphas in past studies have averaged .70 for the Abdel-Halim (1978) scale. Researchers might be advised to trade off the overload cost of four additional questions and use the longer Rizzo, House, and Lirtzman (1970) scales for role conflict and ambiguity in order to increase reliability of measurement.

Relationships between role overload, powerlessness, and affective responses to work appear especially worthy of more investigation in educational administration. The vast majority of research involving role perceptions in all vocational areas has concentrated on role conflict and ambiguity rather than other role perceptions. In education only one prior study (Kottkamp & Mansfield, 1985) has investigated the powerlessness-emotional exhaustion relationship, and we know of no prior study that has investigated the role overload-emotional exhaustion relationship. The magnitude of the zero order correlations suggests that further study of these variables should be fruitful in understanding both occupational stress and job satisfaction among educational administrators.

The basic issue of examining leaders for potential relationships between perceived role stressors, affective responses, and behavioral responses with organizational consequences still seems worthy of pursuit. Most existing studies probe role

perception-affective response relationships rather than role perception-behavioral response relationships, and most of the research which does consider behavioral outcomes was conducted with workers at the lower levels of manufacturing firms. It is an important issue to understand the extent to which perceptions of stressors and affective responses to them are related to leadership behavior. Pursuit of this issue will require care in the choice of behaviors to be assessed and creativity in structuring more independence in measurement.

Note

1. This article is based on a revision of a paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, 1984.

References

Abdel-Halim, A. (1978). Employee affective responses to organizational stress: Moderating effects of job characteristics. *Personnel Psychology, 31*, 561-579.

Andrews, J.H.M. (1965). School organizational climate: Some validity studies. *Canadian Education and Research Digest, 5*, 317-334.

Appleberry, J.B., & Hoy, W.K. (1969). The pupil control ideology of professional personnel in "open" and "closed" elementary schools. *Educational Administration Quarterly, 5*, 74-85.

Arnold, H.J. (1982). Moderator variables: A clarification of conceptual, analytic, and psychometric issues. *Organizational Behavior and Human Performance, 29*, 143-174.

Beck, C.L., & Gargiulo, R.M. (1983). Burnout in teachers of retarded and nonretarded children. *Journal of Educational Research, 76*, 169-173.

Beehr, T.A. (1976). Perceived situational moderators of the relationships between subjective role ambiguity and role strain. *Journal of Applied Psychology, 61*, 35-40.

Caldwell, W.E., & Forney, J.W. (1982, March). *The relationship of role conflict and ambiguity and perceived organizational characteristics between superintendents and principals.* Paper presented to the American Educational Research Association, New York.

Caplan, R.D., & Jones, K.W. (1975). The effects of workload, role ambiguity and Type A personality on anxiety, depression and heart rate. *Journal of Applied Psychology, 60*, 713-719.

Cummings, T.G., & Manring, S.L. (1977). The relationship between worker alienation and work-related behavior. *Journal of Vocational Behavior, 10*, 167-179.

Drummond, T.D. (1970). To make a difference in the lives of children. *The National Elementary Principal, 70*, 32.

Erez, M., & Goldstein, J. (1981). Organizational stress in the role of the elementary school principal in Israel. *The Journal of Educational Administration, 19*, 33-43.

Fox, R.S. (1973). *Climate improvement: A challenge to the school administrator.* Bloomington, IN: CFK Ltd.

French, J.R.P., & Caplan, R.D. (1973). Organizational stress and individual strain. In A.J. Marrow (Ed.), *The failure of success.* New York: AMACOM.

Freudenberger, H.J. (1977). Burnout: Occupational hazard for the child care worker. *Child Care Quarterly, 6*, 90-99.

Friesen, D., Holdaway, E.D., & Rice, A.W. (1984). Factors contributing to the job satisfaction of school principals. *The Alberta Journal of Educational Research, 30*, 157-170.

Gilbert, L.A., & Mangelsdorff, D. (1979). Influence of perceptions of personal control on reactions to stressful events. *Journal of Counselling Psychology, 26*, 473-480.

Gmelch, W., Koch, J., Swent, B., & Tung, R. (1982, March). *What stresses school administrators and how they cope*. Paper presented to the American Educational Research Association, New York.

Gould, R. (1972). The phases of life: A study in developmental psychology. *The American Journal of Psychiatry, 129*, 521-531.

Halpin, A.W., & Croft, D.B. (1962, August). *The organizational climate of schools* (Contract No. SAE 54-8639). Washington, DC: United States Office of Education.

Hartley, M.C., & Hoy, W.K. (1972). Openness of school climate and alienation of high school students. *California Journal of Educational Research, 23*, 17-24.

Hiebert, B. (1985). *Stress and teachers: The Canadian scene*. Toronto: Canadian Education Association.

Hoy, W.K., & Henderson, J.E. (1983). Principal authenticity, school climate and pupil control ideology. *The Alberta Journal of Educational Research, 29*, 123-130.

Indik, V., Seashore, S.E., & Slesinger, J. (1964). Demographic correlates of psychological strain. *Journal of Abnormal and Social Psychology, 69*, 26-38.

Iwanicki, E., & Schwab, R. (1981). A cross validation study of the Maslach Burnout Inventory. *Educational and Psychological Measurement, 41*, 1167-1174.

Jankovic, M. (1983). *Factors associated with school principals' experiences of work related stress*. Unpublished doctoral dissertation, University of Alberta.

Kahn, R., Wolfe, D.M., Quinn, R.P., & Snoek, J.D. (1964). *Organizational stress: Studies in role conflict and ambiguity*. New York: Wiley.

Kanner, L. (1974). *Machiavellianism and the secondary schools: Teacher-principal relations*. Unpublished doctoral dissertation, Rutgers University.

Kornhauser, A. (1965). *Mental health of the industrial worker: A Detroit study*. New York: Wiley.

Kottkamp, R.B., & Helfman, H.L. (1985). *Role perceptions and outcomes: Moderating effects of dogmatism for multidisciplinary team administrators*. Manuscript submitted for publication.

Kottkamp, R.B., & Mansfield, J.R. (1985). Role conflict, role ambiguity, powerlessness and burnout among high school supervisors. *Journal of Research and Development in Education, 18*(4), 29-38.

Kottkamp, R.B., Mulhern, J.A., & Hoy, W.K. (1985). *Secondary school climate: A revision of the OCDQ*. Manuscript submitted for publication.

Kraut, A.I. (1965). *A study of role conflicts and their relationship to job satisfaction, tension, and performance*. Unpublished doctoral dissertation, University of Michigan.

Levinson, D.J. (1974). The psychosocial development of men in early adulthood and the mid-life transition. In D. Ricks (Ed.), *Life history research in psychology*. Minneapolis: University of Minnesota Press.

Lidz, T. (1980). Phases of adult life: An overview. In W.H. Norman & T.J. Scaramella (Eds.), *Mid-life: Developmental and clinical issues*. New York: Brunner/Mazel.

Lusthnus, C.S. (1975). *Job related tension in Quebec school administrators*. Paper presented to the CESEA Conference, Edmonton.

MacPherson, M.A. (1985). Burnout and the school principal. *The Canadian Administrator, 25*(1), 1-4.

Margolis, B.K., Kroes, W.H., & Quinn, R.P. (1974). Job stress: An unlisted occupational hazard. *Journal of Occupational Medicine, 16*, 659-661.

Maslach, C., & Jackson, S. (1978, April). *A scale measure to assess experienced burnout: The Maslach Burnout Inventory*. Paper presented to the Western Psychological Association, San Francisco.

Maslach, C., & Jackson, S. (1979). *The measure of experienced burnout*. Unpublished manuscript, University of California, Berkeley.

McGrath, J.E. (Ed.). (1970). *Social and psychological factors in stress.* New York: Holt, Rinehart & Winston.

Mirvis, P.H., & Lawler, E.E., III. (1977). Measuring the financial impact of employee attitudes. *Journal of Applied Psychology, 62,* 1-8.

Miskel, C., Fevurly, R., & Stewart, J.W. (1979). Organizational structure and processes, perceived school effectiveness, loyalty and job satisfaction. *Educational Administration Quarterly, 15,* 97-118.

Mueller, E. (1965). *Workload of university professors.* Unpublished doctoral dissertation, University of Michigan.

Mullins, J. (1976). *Analysis and synthesis of research utilizing the Organizational Climate Description Questionnaire: Organizations other than elementary schools.* Unpublished doctoral dissertation, University of Georgia.

Pincherle, G. (1972). Fitness for work. *Proceedings of the Royal Society of Medicine, 4,* 321-324.

Porter, L.W., & Steers, R.M. (1973). Organizational work and personal factors in employee turnover and absenteeism. *Psychological Bulletin, 80,* 151-176.

Rizzo, J.R., House, R.J., & Lirtzman, S.I. (1970). Role conflict and ambiguity in complex organizations. *Administrative Science Quarterly, 15,* 140-169.

Schwab, R.L., & Iwanicki, E.F. (1982). Perceived role conflict, role ambiguity, and teacher burnout. *Educational Administration Quarterly, 18*(1), 60-74.

Schwandt, D.R. (1978). *Analysis of school organizational climate research 1962-1977: Toward construct clarification.* Unpublished doctoral dissertation, Wayne State University.

Seeman, M. (1959). On the meaning of alienation. *American Sociological Review, 24,* 783-791.

Shepard, J. (1972). Alienation as a process: Work as a case study in point. *Sociological Quarterly, 13,* 161-173.

Swent, B., & Gmelch, W.H. (1977). *Stress at the desk and how to cope creatively.* Eugene: Oregon School Study Council.

Syrotuik, J., & D'Arcy, C. (1983). *Saskatchewan teacher stress study: Report to respondents.* Saskatoon, SK: Applied Research Unit, Saskatchewan Health Department.

Thomas, A.R. (1976). The organizational climate of schools. *International Review of Education, 22,* 441-463.

Travlos, A.L. (1984). *Selected job stressors, job satisfaction, emotional exhaustion, and thrust behavior of the principal.* Unpublished doctoral dissertation, Rutgers University.

Tung, R., & Koch, J. (1980). School administrators: Sources of stress and ways of coping with it. In C.L. Cooper & J. Marshall (Eds.), *White collar and professional stress.* New York: Wiley.

Vroom, V.H. (1964). *Work and motivation.* New York: Wiley.

Waters, L.K., & Roach, D. (1971). Relationship between job attitudes and two forms of withdrawal from the work situation. *Journal of Applied Psychology, 55*(1), 92-94.

Zedeck, A. (1971). Problems with the use of "moderator" variables. *Psychological Bulletin, 76,* 295-319.

Zeilinski, A.E., & Hoy, W.K. (1983). Isolation and alienation in elementary schools. *Educational Administration Quarterly, 19*(2), 27-45.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September, 1986, 249-264.*

PERSPECTIVES

Making Research Matter

Edward A. Holdaway
University of Alberta

In many countries researchers have expressed concern over the perceived lack of impact of their findings upon social policies. The relationship between research and policy has been the subject of several recent major conferences and publications. Many explanations have been presented together with related suggestions for increasing impact. However, the complexity of the relationship and the need to distinguish between different types of research and different types and levels of policy have often been overlooked. Indirect long-term impact is probably the most important research contribution. The ways in which social science research contributes to knowledge require more examination.

Since the mid-1960s, and especially during the last decade, many writers have focused their attention on the impact of research upon social policy. This arose partly from a concern that a great deal of research, involving considerable effort and expense, appeared to have little effect upon policy, and partly from the need to show practical benefits from research in order to defend research expenditures. The purposes of this article are to present evidence that this concern persists internationally, to reassess the reasons why direct impact continues to be less than expected, to suggest an attitude which independent researchers should adopt, and to propose some approaches which could increase the impact of policy-relevant research. Selected recent literature, personal experiences, and interviews conducted by the author in several countries are used as sources.

Throughout this article, "research" is viewed as "systematic critical investigation directed toward increasing the general body of knowledge or understanding of the subject studied" (Radford & McDowell, 1973, p. 11). As Nisbet and Broadfoot (1980, p. 42) have emphasized, research obviously includes a very wide range of activities and methodologies. Some writers, including the author, categorize

Dr. Holdaway is a Professor of Educational Administration and Director of Institutional Research and Planning whose current research interests focus on job satisfaction.

"evaluation" as a type of research, whereas others classify evaluation and research as separate activities. "Policy" is viewed as either official statements or unwritten understandings which direct or guide behavior.

Persistence of Widespread Concern and Interest

Evidence of increasing attention to the impact of research upon policy is provided by the following publications: (a) books prepared by Nisbet, Megarry, and Nisbet (1985), Husén and Kogan (1984), Bulmer (1982), Adams, Smelser, and Treiman (1982), and Nisbet and Broadfoot (1980); (b) special articles in *Administrative Science Quarterly* in 1982; and (c) the 1981 Special Issue of the *International Review of Education*. Note must also be taken of the important works of Coleman (1972), Taylor (1973), Caplan, Morrison, and Stambaugh (1975), Weiss (1977), Suppes (1978), and Mitchell (1980). The topic has received attention from various organizations including the US National Academy of Sciences, the Social Science Research Council of Great Britain (recently renamed the Economic and Social Research Council), the Australian Council for Educational Research, the Foundation for Educational Research in the Netherlands (SVO), the Council of Europe, and the US National Research Council. As Lindblom and Cohen (1979, p. 1) have stated, "In public policy making, many suppliers and users of social research are dissatisfied, the former because they are not listened to, the latter because they do not hear much they want to listen to." Information and opinions garnered from Canadian and other sources attest to the prevalence of this condition.

In a Science Council of Canada study which began in 1978, Karen Fish (1983) conducted 40 interviews with Members of Parliament (Ottawa), Senators, parliamentary staff members, and members of scientific organizations. She concluded that "scientific and technological factors are peripheral to the economic, social and political concerns determining parliamentary decisions" (p. 9). This study was undertaken "to assess government's ability to respond to 'the complex and divisive issues that arise from the interaction of science and human values,'," (p. 5). The associated difficulties in policy making are exemplified by the commonly expressed views that more research should be conducted before action is taken to reduce the environmental effects of acid rain. Another important, recent example at the federal level in Canada is provided by the decisions of the Department of Justice and Transport Canada to obtain "scientifically accurate data that will help them make informed decisions about the areas for which they have responsibility," including action to reduce the severe consequences of accidents caused by drinking drivers (Supply and Services, 1985, p. 11).

The Canadian social science community has been actively involved in the research-policy debate as evidenced by articles and reports by Breton (1977), Deiseach (1978), the Social Science Federation of Canada (1983), and Coleman and LaRocque (1983). These have mainly addressed the contributions of research and the reasons why such contributions are often less than expected. On the other hand, questions are frequently raised about the apparent lack of relevance of much Canadian social science research. Andre Fortier (1978, pp. 11-12), former SSHRCC President, observed that "Many MPs are worried about taxpayers' money being spent on projects which do not have any bearing on Canadian problems. Journalists increase this concern by writing sensational articles." This perception was shared by John Trent (1982, p. 2), formerly Executive Director of the Social Science

Federation of Canada:

> Few Canadian social scientists recognize that we are also criticized for being esoteric and irrelevant [and] for not being a source of creative ideas in these times of socioeconomic crisis. . . . it is clear that given our current level of personnel and resources, governments expect social scientists to be seen to be making greater contributions to national, applied, interdisciplinary, problem-oriented research.

In Britain, Nisbet and Broadfoot (1980) produced one of the most comprehensive treatises on this subject. They identified the failure of "the greatly increased investment in educational research and curriculum development" to meet "raised hopes and expectations" as a major source of concern in the UK over the research-policy gap. They also felt that "growing sophistication and self-awareness among educational researchers led to aspirations to hold an established place in the educational system" which were apparently not realized. Evidence of continued interest in this area was provided by the recent meeting of a group of social scientists with MPs in the House of Commons to explore the possibility that they could play a greater part in British policy making (Wotjas, 1984). Levy and Healey (1984), however, cautioned that a danger exists in overselling the contribution of social science to the management of change in the current technological revolution. In a related comment, Hamilton (1984) stated that greater coordination was needed if educational research wanted to overcome its fragmented impact upon British educational policy.

Similar views have been expressed in the United States. For example, in the 1978-1979 Report of the US National Council on Educational Research, the Chairman, John Corbally (1980, p. 1), stated that the brunt of concerns raised by members of Congress, their staff, and interested groups after the first year of existence of the National Institute of Education was, "You spent all of that money (around $140 million in the first 'golden year') and the schools still aren't any better." Also, Caplan (1982, p. 36), after conducting a study in which over 200 upper-level federal executives were carefully questioned, concluded that formal or informal contact between social scientists and senior executives was rare. In order to understand more clearly the impact of social science research findings on policy and so focus attention upon the formulation of educational policies by US state legislatures, the National Institute of Education funded a study by Mitchell (1981) who collected data from 160 "key legislative policy influentials" in Arizona, California, and Oregon. Mitchell's (1981, p. 299) major conclusion was that "scientific expertise makes important contributions to legislative policy only if policymakers are oriented toward a reliance on expert authority and authentic scientific expertise is available to them." In connection with this type of activity, Holden (1981, p. 1398) observed that a lot of people are making lists to demonstrate all the ways that social and behavioral research has contributed to society.

The same types of concerns have been expressed in New Zealand. For example, Geraldine McDonald, in her 1980 NZARE Presidential Address, remarked that:

> Nevertheless there appears to be a strong urge both in New Zealand and in countries other than our own to try to bring research and policy into a closer relationship, to bring relevant research to the attention of those who make decisions, to understand the problems of decision makers, and to analyze the processes by which policy is determined.

Since 1977, the Australian Council for Educational Research (ACER) (1980, p. 30) has undertaken a research program "specifically designed to inform social policy decisions." This program consists of three groups of studies: (a) evaluation of selected policies and programs; (b) transition from school to adult life; and

(c) allocation of staff and resources to schools. ACER feels that such research contributes evidence to assist decision making as well as analyzing and understanding the issues being considered. Also, Dunn (1981) cited the opinion of the former Chairman of the Australian Schools Commission that research has had little impact upon the work of the Commission should have had more. Further, some valuable insights into the importance in policy making of politicians and state directors-general of education in Australia have been provided in the eight monographs of the University of Melbourne project entitled *The Education Policy Process at State Level: An Australian-United States Comparative Study* (Project Directors, G.S. Harman & F.M. Wirt). These monographs include many examples of policies which have been adopted by premiers, governments, and directors-general, following differing amounts of public involvement.

Explanations

Before presenting the reasons why research commonly has had less impact upon policy than is expected, three major points must be emphasized. First, even the most cynical must accept that some research activities have had an obvious and direct impact upon the formation of policies. The National Academy of Sciences 1977 Report assessed that systematic, disciplined inquiry has helped in the transformation of education over the century, citing examples such as greater access to education and the development of more humane attitudes. That report also stated that people throughout the world see and talk about reality differently as a result of what Freud, Dewey, Thorndike, Skinner, Piaget, and Mead have written and said (Kiesler & Turner, 1977, pp. 11-12). Research findings have also been influential in (a) the comprehensivization of Swedish secondary schools (Husén, 1978, p. 545); (b) positive discrimination by use of additional funding for "deprived areas" identified in the Plowden Report (Bulmer, 1982, p. 41); and (c) the decision to delete French from British primary schools (personal communication Yates, 1982). Second, we must recognize that some educational policies, such as the provision of "a balance of social, physical, emotional, and moral development" for students (Education Department of South Australia, 1981), have little obvious research base. Third, the relationship between research and policy must be viewed as more complex than was originally supposed (Foster, 1979, p. 93). The complexity of the research-policy relationship was referred to by Caplan (1982, pp. 32-33) in this way:

> As a result of recent empirically-based studies, it has become evident that even with the increased production of policy-relevant knowledge and improved technological procedures for transfer and dissemination, the frequency of use and impact of knowledge has not increased substantively. Simply because information is timely, relevant, objective, and given to the right people in usable form, its use has not been guaranteed. . . . Obviously, utilization is not a simple process.

This is compounded by the complexity of the processes involved in gaining knowledge and forming policy. Weiss (1982) coined the terms "knowledge creep" and "decision accretion" to indicate that obtaining increases in knowledge and arriving at decisions are both slow, gradual processes. Warren Piper (1980) considered that governments may be well informed by social science, but are weak in producing workable solutions. Mitchell (1981, p. 271) added further insight into the complexity issue by presenting these conclusions based on six empirical survey studies of social science utilization: (a) the concept of "social science" is complicated and difficult to interpret; and (b) "utilization" has several different dimensions as variations occur in the frequency and the extent of use of the findings.

Many reasons have been advanced in the literature to explain why educational research activities have not obviously affected educational policy in a substantial and direct manner. A comprehensive categorization of these reasons has been presented by Holdaway (1982). Many relate to differences in orientation, knowledge, and time-frames of policy makers and researchers. Henkel and Kogan (1981, p. 39), Nuthall (1975, pp. 4-5), Caplan (1982, p. 33), and Davey (1982, p. 1) have all affirmed that policy makers prefer to rely on their own judgment. In Caplan's words (1982, p. 33):

> Bureaucratic, ethical, attitudinal, and social considerations take precedence over the value of information in its own right, and that is probably the way it should be. After all, social problem-solving is a political process and it is neither desired nor expected that it should be otherwise.

Crossland (Boyle, Crossland, & Kogan, 1971, p. 190) lent support to Caplan's conclusion in describing Britain's change to comprehensive schools as being influenced by basic value judgments about equity, equal opportunity, and social division, as well as views about education. Another aspect was noted by Mitchell (1981, p. 299) who found that "scientific expertise makes important contributions to legislative policy only if policymakers are oriented toward a reliance on expert authority and authentic scientific expertise is made available to them."

Also identified as a reason for less than expected impact are the reports written by social scientists. For example, Crelinsten (1984, p. 2) considered that, in attempting to be scientific, social scientists have given their research a bad name by writing obscure, pedantic, and jargon-filled articles. Davey (1982, p. 1) concluded that politicians commonly view research as biased, outdated, long-winded, academic, and impractical. Researchers are also criticized because they are used to making different explanations, which is viewed as hedging by laymen (Paulston, 1982, p. 9), because they fail to deliver the findings relevant to crucial decisions (Baumgart & Low, 1979, p. 77), and because their studies lead to conflicting results (Averch, Carroll, Donaldson, Klesling, & Pincus, 1972, p. 2). Steinle (1981, p. 606) added another significant concern—administrators do not have the time to read all of the relevant research.

Further, as Wirt and Mitchell (1982, p. 2) concluded, knowledge will be utilized only if it enhances political influence, simplifies policy problems, or assists in identifying effective solutions. Stuart Smith (1984), President of the Science Council of Canada, considered that social scientists have not been sufficiently aggressive in demonstrating applications of their findings to policy. Another major reason relates to "time": Whereas politicians work within a very short time frame, researchers can rarely conduct the necessary studies in the time available. Tizard (personal communication, 1982) observed that politicians who set up a project are often gone by the time the project is completed. Levy (personal communication, 1982) has also noted that many disciplines, such as economic history, are always closely related to policy, whereas others such as archaeology are not.

Another explanatory thrust has been provided by Lindblom and Cohen (1979, pp. 10-29) in their discussion of "ordinary knowledge" and of social learning and interactions as complementary alternatives to professional social inquiry (PSI). In their view, social scientists underestimate society's use of "new ordinary knowledge" which comes from "common sense, causal empiricism, or thoughtful speculation and analysis" (p. 12). PSI is seen to add little real new knowledge and much of it is ordinary knowledge, such as "television absorbs much of children's time." Most PSI practitioners "rely heavily on the same ordinary techniques of speculation,

definition, conceptualization, hypothesis formulation, and verification as are practiced by persons who are not social scientists or professional investigators of any kind" (p. 15).

These views of Lindblom and Cohen about the sources of "ordinary knowledge" were not fully supported by Weiss (1982, p. 623) and Rule (1982, p. 4) who argued that, because social science knowledge and concepts become assimilated with everyday knowledge and language, policy makers commonly are unaware that they are using research findings. Prewitt (1981) provided examples of social science concepts common to our vocabulary including human capital, identity crisis, span of control, price elasticity, and acculturation.

Lindblom and Cohen (1979) also claimed that policy formation can be heavily influenced by social learning; for example, an energy policy may be more acceptable after people have waited in line for gasoline. They also concluded that "we do not well understand that interaction is a method of problem solving" (p. 20). Interactive problem solving was defined by them as "to undertake or stimulate action—usually interaction—so that the preferred outcome comes about without anyone's having analyzed the given problem or having achieved an analyzed solution to it" (p. 20). They classify voting, including legislative voting, as interaction.

Obviously, many explanations have been proposed for the perceived lack of impact. Probably all of these explanations have some validity in the case of individual policies. The caution of Lindblom and Cohen (1979) concerning the marginal growth of knowledge through professional social inquiry, is important, even if we can identify some research results, such as Herzberg's well known two-factor theory of motivation, which were not part of ordinary knowledge. However, a need still exists for research findings to assist policy making. As Inger Marklund (1981, p. 14) has stated: "Perhaps the greatest paradox of all is that, at the same time as more and more is being banked on R & D at the political level, confidence in the problem solving capacity of R & D is declining." The existence of a "chicken and egg" situation in this regard was noted by Nisbet and Broadfoot (1980, p. 45): Research lacks impact because it is underdeveloped, but its lack of direct use contributes to continuing underdevelopment.

Types of Impact

Further explanations can be derived from a consideration of the three different types of impact identified by Nisbet and Broadfoot (1980, pp. 20-21):

1. *Immediate, direct impact.* This is perceived to be rare except when topics are noncontroversial and values are not challenged. Taylor (1981, p. 190) considered that research which has the strongest short-term impact may not have the greatest long-term influence, but its publicity could generate further studies which advance knowledge and improve the rationality of decision making.
2. *Intermediate impact.* This occurs with evaluation and developmental work, which is seldom published, and which obtains its original impetus from research.
3. *Long-term, indirect impact.* Nisbet and Broadfoot saw this impact as that which "influences the climate of opinion" and concluded that it "is increasingly recognized in the literature as the most important and far-reaching." Authors who supported this view include Angus (1981, p. 575), Getzels (1978,

pp. 487-488), Nuthall (1975, pp. 4-5), Powell (1980, p. 2), Power (1981, p. 14), Wagenaar, Kallen, and Kosse (1982, p. 8), and Weiss (1982, p. 621). They perceived that basic research is more influential in the long run than is research focused upon a particular practical problem because it has the capacity to provide different conceptualizations, to reformulate problems, to furnish new insights, to sharpen perceptions, and to stimulate discussion and questioning. Adams, Smelser, and Treiman (1982, p. 102) drew this conclusion:

> The power of basic research to improve and enrich our lives grows out of the mutual reinforcement and synergism of many interlocking ideas, findings, and practical outcomes. It cannot be understood and properly utilized if we concentrate instead on isolated, product-centered outcomes.

Another indirect aspect was raised by Breton (1977, p. 45) who reasoned that policy research will have an impact upon policy makers if it affects their cognitive system, that is, the system of ideas, postulates, and perceptions of reality. Further, Beyer and Trice (1982, p. 605) contended that the results of applied research are "often not generalizable to other situations."

Bulmer (1982) has provided two models which place many of the foregoing explanations in perspective. The *engineering model*, which is called the physical science model by some authors, presumes that scientific findings will lead to desired end-states agreed upon by both researchers and decision-makers, whereas the *enlightenment model* presumes that one purpose of research is to provide background information which can alter the conceptions of policy makers. This latter model of indirect impact obviously is favored by most writers who have examined this field. Another useful categorization was provided by the National Academy of Sciences 1977 report entitled *Fundamental Research and the Process of Education* (Kiesler & Turner, 1977, pp. 8-11) which proposed these four spheres of influence of research:

1. *A view of reality*—research influences how people view the world, the languages that they use, and the values they attach to their observations;
2. *A vision of the achievable*—research influences our expectations and understandings, for example, what can be achieved by handicapped children and what limitations are imposed on underprivileged groups;
3. *Know-how*—research may lend support to our current practices, and we should not always consider that research should lead to improvement; and
4. *A commitment to act*—research can provide data to support the view that an activity (e.g., teaching) is worthwhile, as well as assisting in improvement, altering conceptions, and deepening work commitment.

A different set of models of the process by which research enters social problem-solving was presented by Weiss (1977):

1. *Decision-driven model*—research responds to a problem identified by decision makers;
2. *Knowledge-driven model*—research presents an opportunity that can be capitalized upon by decision makers; and
3. *Interactive model*—research is part of a complex search for knowledge from a variety of sources.

This third model is the one favored by Lindblom and Cohen (1979) as described earlier. Weiss's first model relates to the work of commissions or committees of inquiry, which many authors see as having influenced policy far more than have individual research projects. To a large extent this phenomenon has resulted because

governments established commissions to investigate and report upon areas in which governments already have decided that policy changes are desirable. Bulmer (1982, p. 99) claimed that "Commissions may also act as a device to enable governments to do what they want to do anyway, sometimes clothing it in the legitimacy provided by research."

After reviewing the impact of Canadian commissions investigating postsecondary education (PSE) from 1965-1980, Ahamad (1980) concluded that some were more influential than others primarily because of timing, membership, presentation, and marketing of the report and implementation plans. Further, Sheffield (1982, p. 51) observed that

> Even though some provincial inquiries have not had their recommendations accepted, it must be said that such projects make use of existing research results and create more. As a consequence, the stock of knowledge about PSE is increased and enriches the resources available to future researchers and decision-makers.

James Coleman (1984, p. 132) has made an interesting, related observation as a result of experiences with his US Equal Education Opportunity Report: Some social policy research is used in unintended ways. For example, research associated with his report was never used in formation of civil rights education policy, but it was considered in resolution of conflicts arising from local school desegregation.

Lord Rothschild (1982, pp. 18-21) supported the need for research when commissions are considering policy options. In his review of the British Social Science Research Council, he observed that failures in public policy and planning since 1945, such as the building of high-rise public housing apartment blocks, are "in part attributable to the lack of adequate enquiry into social facts." He felt that slum clearance and the building of new towns required "organized enquiries of the kind that social scientists are trained to undertake," and concluded that

> there are, and will continue to be policy issues to which empirical research, requiring high academic standards in the enquiry, are directly relevant. Such research is, at the very least, a check on the credibility of entrenched common sense beliefs, which may not correspond with the facts.

This view calls into question the appropriateness for policy making of "ordinary knowledge" as described by Lindblom and Cohen. Of course, the concern that social scientists feel about the impact that their work has or could have, as in research on the areas identified by Lord Rothschild, would depend upon their perceptions of their role. In this regard, Dunn (1981, pp. 519-525) provided a useful distinction between "technocratic guidance" and "technocratic counsel." The technocratic guidance perspective emphasizes a power shift to policy analysts and away from policy makers, generated by "changes in the structure of government and the nature and scope of social problems" and the view that "policy-relevant knowledge is an increasingly scarce resource." The technocratic counsel perspective emphasizes that an increase in policy-analysis activities enhances the power of dominant groups through legitimization of their policy decisions. Dunn concluded that neither perspective is accurate, as the technocratic guidance approach exaggerates the power and influence of analysts while the technocratic counsel approach exaggerates the symbolic importance of legitimization. However, these two perspectives do produce ideal types which aid analysis and understanding.

Suggestions

Those who have examined the impact of research upon policy have made many suggestions for increasing this impact. These suggestions can be categorized under the headings of improving research and improving dissemination and interaction.

Improving Research

Suggestions for improving research include the use of more thematic research which is focused upon particular topics: This would allow greater concentration of research and the obtaining of more information relevant to an area of policy interest. This approach has been employed in recent years by the Social Sciences and Humanities Research Council of Canada, although it has been criticized by the advocates of independent, scholar-generated research. Others recommend more triangulation—the use of different methods simultaneously or successively, in an attempt to obtain greater reliability and thereby more faith in the results. Further suggestions include narrower delineation of research problems, more frequent use of theoretical frameworks, more emphasis on illuminating case studies, and even greater concentration on basic research which many see as having the most long-term utility.

Another set of suggestions relates to making research more directly "policy-relevant." Coleman and LaRocque (1983, pp. 247-248) have discussed this matter as follows:

> In order for research findings to be utilized . . . research should be addressed specifically to policy-makers, should investigate problems which are of concern to them, and should be designed to provide data and interpretations which have predictive utility concerning the probable consequences of policy options. Furthermore PRR [policy-relevant research] should concern itself with those policy options which are important to policy-makers, because they are normatively acceptable and have some face validity as solutions to the perceived problems. . . . Additionally, policy-relevant research should be concerned with utility at every stage of the research, rather than addressing this as a postscript on policy implications, and hence should:
>
> a. focus on variables manipulable by policy makers . . . and
>
> b. be sensitive to the political environment and to uses likely to be made of the research findings.

McDonald (1976, pp. 1-2) also advocated that consideration of policy alternatives should pervade the work of educational researchers and that "the policy alternatives available to the policy makers should be carefully laid out as a first step in planning research."

Improving Dissemination and Interaction

Suggestions about communication include the development of mutual confidence by researchers and policy makers, greater collaboration and cooperation, sharing of orientations, better understanding of the research-policy relationship (as in Mitchell's work), improved dissemination techniques, and greater monitoring of research findings having potential relevance to areas of policy interest. Dube (1982, p. 7) has succinctly addressed several of these matters:

> It is necessary to establish better rapport and communication between the social scientists and the policy makers. Often the technical jargon of the social scientists puts off the practical minded administrators and policy makers. The social sciences findings have to

> be communicated to them in a different idiom. And bridges have to be built between the two so that their uneasy relations are replaced by more harmonious and interpenetrating work relationships.

However, Weiss (1982, p. 637) has cautioned that "adherence to all the traditional strictures," including "jargon-free communication," has only a marginal effect upon the actual application of results.

Further, Mitchell (1981, p. 153) recommended that universities should develop training programs for legislative staff to improve their capabilities in location and interpretation of policy-relevant social science findings. At the same time, he cautioned that "interpretation of social science is every bit as complex and professional a task as is its creation." Nisbet and Broadfoot (1980, p. 12) emphasized that change is not effected merely as a result of telling educators about research findings. Another dimension was highlighted by Lindblom and Cohen (1979, pp. 2-3): The traditional parochial views about inadequate utilization (e.g., unsatisfactory personal relations between researcher and client) ignore the wider issue that man is a social problem-solver who does not rely to a great extent on research findings.

Analysis

The wide variety of views on many aspects of the impact of research upon social policy, which Lindblom and Cohen (1979, p. 1) rightly called "a river of mixed diagnosis and prescription," require some analysis and synthesis. Such an activity necessarily involves subjective assessment and selection based partly upon personal experience, preference, and force of argument. A summary of the major points advanced earlier in the article is first presented. Second, attention is given to aspects which were missed or unemphasized. Third, some areas of disagreement are reported. Fourth, some suggestions are presented for researchers and those interested in the topic of research impact.

Summary

The major points or lines of argument developed thus far are as follows:

1. Many writers have paid attention, over the last two decades, to the apparent lack of impact of research findings in social science, including education, upon policy formation;
2. This interest is persisting on an international basis;
3. The perceived lack of impact is of concern to both researchers and policy makers;
4. Policy usually results from social, economic, and political forces and from personal sentiment rather than from research findings;
5. Some obvious cases of direct impact have been identified, especially when the research has been related to commissions of inquiry;
6. Many reasons and suggestions have been advanced for explaining and reducing the lack of direct impact;
7. The greatest contribution of research is through indirect influence by clarifying issues, providing data, sensitizing policy makers, and providing conceptual analyses;
8. At the same time that the contribution of social science research towards policy problems and the solution of social problems is being questioned, policy makers in a wide variety of arenas are requesting that more research be

conducted to assist them;

9. The relationship between research and policy is complex, resulting in part from different types of research, different types of policy, and different types of applications of research findings; and
10. Much of the knowledge used by policy-makers comes from sources other than social research.

Underemphasis

Some aspects of the phenomenon appear to have been either omitted or somewhat underemphasized in the cited literature. First, the term "relationship" is probably better than "contribution" or "impact" to emphasize that the research-policy linkage is interactive rather than unidirectional. To illustrate, Houston (personal communication, 1982) has pointed out that policy and policy-making have had a profound effect upon the research enterprise in many western countries. In particular, he cited the work of Federal Government commissions in Australia in initiating a great deal of research activity. Second, few writers have considered the existence of policies at different levels: in education, for example, we can have policies at the federal, state/provincial, district, and school levels. Mitchell's study of three US legislatures, although highly important, focused only on the impact of research on state policies. Similar research on impact at other political and administrative levels is warranted. Third, the epistemological concerns raised by Lindblom and Cohen, especially with respect to the relationship of ordinary knowledge to social science knowledge, have not been addressed by most authors. Although Lindblom and Cohen may have somewhat overstated the minimal contribution of our research to "new" knowledge, their view deserves serious consideration by all educational researchers. To what extent are we producing "new" knowledge as compared with knowledge which is already in the public domain? Fourth, the difficulty of ascribing causality to actions has received little recognition, except through discussion of the major indirect effects. Even identifying the influence of relevant research findings upon a particular policy decision is a very difficult exercise. A fifth point, made by only a few writers, relates to the distinction between scholar-generated research and "in-house" research, such as university/college institutional research or directed/contract research performed for a commission or committee of inquiry or government agency. The findings of such research, clearly targeted upon a particular problem area, are far more likely to have substantial impact. Hence, those reflecting about the research-impact relationship should clearly distinguish between these two major types of research activity as they have different purposes and consequently different associated expectations. Those analyzing in-house research activity, which may require researchers to act as policy advisors, should take cognizance of Dunn's polar extremes of technocratic counsel and technocratic guidance.

Sixth, although improved interaction between researchers and policy makers (both elected and appointed officials) is frequently encouraged, insufficient emphasis has been placed upon analyzing the types of interaction leading to policy that Lindblom and Cohen have discussed. Their view is supported by Warren Piper (1980) who argued that we need to develop models of human interaction involving the application of occupational and social psychology to educational management. The seventh point to be made in this section relates to policy studies. Laswell's distinction between studies *of* policy and studies *for* policy was deemed useful by

Downey (1984, p. 1) who saw the first type as relatively unproblematic and the second type as somewhat controversial. This again highlights the importance of identifying different types of research when generalizations about impact are being made. Downey (1984, p. 1) also proposed that the disagreements concerning the use of research, different sources of knowledge, and the roles of analysis and interaction in policy development (Lindblom and Cohen, 1979) "ought to become the bases of continuing dialogue between those who do policy research and those who shape policies for education." Lastly, only a few writers refer directly to the existence of different types of impact, such as the classification presented by the National Academy of Sciences (Kiesler & Turner, 1977).

Disagreements

Probably the major disagreement observed in the literature relates to the issue of whether impact can be increased, and if so, by what means. Many advocate that the impact should be increased and that this can be achieved by improving existing or using new procedures. However, others consider that we should not endeavor to increase direct impact, because the greatest impact is indirect, occurring through a variety of measures which alter perceptions.

Suggestions

The study of the relationship between research and policy obviously is wide-ranging, complex, and important, and warrants our continuing attention. My assessment is that virtually all of the concerns, explanations, and suggestions voiced by others have some validity. Even if Lindblom and Cohen's view of social knowledge generation is accepted, there surely is room for more interaction between researchers and policy makers, especially with respect to policy-relevant research. No one view should prevail, to the exclusion of others, in the debate about how to make research matter.

Moreover, educational researchers should neither confuse nor combine basic research and policy-oriented research. Basic research can have a long-term effect on policy (and practice), but immediate, direct linkage should not be expected. A great deal of educational research conducted by university staff and students has the primary purpose of increasing our knowledge and understanding. To expect that such research will necessarily have policy implications does a disservice to the research and policy communities as well as to the activity of educational research. The suggestions made by Coleman and LaRocque (1983) are pertinent here. All too frequently both academic staff and graduate students in education conduct theoretical research, whose aim is to advance knowledge and understanding, but in the concluding sections feel obliged to discuss the implications of their findings for policy without having paid any reference to such implications throughout the design, data collection, or data analysis phases. Accordingly, the relationship between research and policy should receive more attention in graduate courses on research methodology.

A fertile area for further examination is the means by which social science findings relate to our knowledge, through assimilation which changes vocabulary and attitudes, through finding empirical support for "ordinary knowledge," or occasionally through the relatively direct impact of policy-oriented research. In view of the implications of policies for practice, and the various concerns described in this article, such continuing examination of the research-policy relationship is easily justified.

References

Adams, R.M., Smelser, N.J., & Treiman, D.J. (Eds.). (1982). *Behavioral and social sciences research: A national resource.* Washington, DC: National Academy Press.

Ahamad, W. (1980). *The undigested studies of the seventies.* Paper presented at the Annual Conference of the Canadian Society for the Study of Higher Education, Montreal.

Angus, M. (1981). Making policy research in education more directly useful. *Proceedings of the Annual Conference of the Australian Association for Research in Education* (pp. 575-590). Adelaide: South Australian Government Printer.

Australian Council for Educational Research. (1980). *Fiftieth annual report 1979-80.* Hawthorn: Australian Council for Educational Research.

Averch, H.A., Carroll, S.J., Donaldson, T.S., Klesling, H.J., & Pincus, J. (1972). *How effective is schooling?* Santa Monica, CA: Rand.

Baumgart, N.L., & Low, B.C. (1979). The establishment of dialogue between researcher and policy maker. In J.S. Shellard (Ed.), *Educational research for policy making in Australia* (pp. 77-92). Hawthorn, Victoria: Australian Council for Educational Research.

Beyer, M., & Trice, H.M. (1982). The utilization process: A conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*, 27(4), 591-622.

Boyle, E., Crossland, A., & Kogan, M. (1971). *The politics of education.* Harmondsworth: Penguin.

Breton, R. (1977). *The Canadian condition: A guide to research in public policy.* Montreal: Institute for Research on Public Policy.

Bulmer, M. (1982). *The uses of social research.* London: Allen and Unwin.

Caplan, N. (1982). Social research and public policy at the national level. In D.B.P. Kallen, G.B. Kosse, H.C. Wagenaar, J.J.J. Kloprogge, & M. Vorbeck (Eds.), *Social science research and public policy-making: A reappraisal* (pp. 32-48). Windsor, Berkshire: NFER-Nelson.

Caplan, N., Morrison, A., & Stambaugh, R. (1975). *The use of social science knowledge in policy decisions at the national level.* Ann Arbor: Institute for Social Research, University of Michigan.

Coleman, J.S. (1972). *Policy research in the social sciences.* Morristown, NJ: General Learning Press.

Coleman, J.S. (1984). Issues in the institutionalization of social policy. In T. Husén & M. Kogan (Eds.), *Educational research and policy: How do they relate?* (pp. 131-141). Oxford: Pergamon.

Coleman, P., & LaRocque, L. (1983). Linking educational research and educational policy via policy-relevant research. *The Alberta Journal of Educational Research, 29*(3), 242-255.

Corbally, J.E. (1980). Report of the Chairman. *Fifth Report of the National Council on Educational Research.* Washington, DC: National Institute of Education.

Crelinsten, J. (1984). Spotlighting the social sciences. *ACCESS, 2*(2), 2 (Association for the Advancement of Science in Canada, Ottawa).

Davey, J.A. (1982). Research and its use. *SSRFC Newsletter, 3*(2), 1-2 (Social Sciences Research Fund Committee, New Zealand).

Deiseach, D. (Ed.). (1978). *Educational research and policy formation.* Toronto: Canadian Education Association.

Downey, L.W. (1984). Perspectives on policy studies. *Policy explorations 1*(1), 1-2 (University of British Columbia).

Dube, S.C. (1982). Crisis and commitment challenges to intellectual craftmanship in the social sciences. *RIHED Bulletin, 9*(3), 4-7 (Regional Institute of Higher Education and Development, Singapore).

Dunn, S.S. (1981). 'The parable of the sower'—its relevance to research in education. *Proceedings of the Annual Conference of the Australian Association for Research in Education* (pp. 519-525). Adelaide: South Australian Government Printer.

Dunn, W.N. (1981). *Public policy analysis: An introduction.* Englewood Cliffs, NJ: Prentice-Hall.

Education Department of South Australia. (1981). *Into the 80s.* Adelaide: Government Printer.

Fish, K. (1983). *Parliamentarians and science.* Ottawa: Science Council of Canada.

Fortier, A. (1978, May). *Communication.* Unpublished paper delivered to First General Assembly of the Social Science Federation of Canada, Ottawa.

Foster, P.J. (1979). World-wide policy-oriented research: Some dilemmas. In J.S. Shellard (Ed.), *Educational research for policy making in Australia* (pp. 93-102). Hawthorn, Victoria: Australian Council for Educational Research.

Getzels, J.W. (1978). Paradigm and practice: On the impact of basic research in education. In P. Suppes (Ed.), *Impact of research on education: Some case studies* (pp. 477-521). Washington, DC: National Academy of Education.

Hamilton, J. (1984, December 7). Report of speech. *Times Higher Education Supplement*, p. 1.

Harman, G.S., & Wirt, F.M. (1981). *The education policy process at state level: An Australia-United States comparative study.* Melbourne: Centre for the Study of Higher Education, University of Melbourne.

Henkel, M., & Kogan, M. (1981). *The DHSS funded research units: The process of review.* London: Department of Government, Brunel University.

Holden, C. (1981, March 27). Dark days for social research. *Science, 211*, 1397-1398.

Holdaway, E.A. (1982). The relationship between research and policy formation in education. In R.F. Broadbent (Ed.), *Education policy making in Australia* (pp. 17-40). Melbourne: Australian College of Education.

Husén, T. (1978). Educational Research and educational reform: A case study of Sweden. In P. Suppes (Ed.), *Impact of research on education: Some case studies* (pp. 523-579). Washington, DC: National Academy of Education.

Husén, T. (1984). Research and policymaking in education: An international perspective. *Educational Researcher, 13*(2), 5-11.

Husén, T., & Kogan, M. (1984). *Educational research and policy: How do they relate?* Oxford: Pergamon.

Kallen, D.B.P., Kosse, G.B., Wagenaar, H.C., Kloprogge, J.J.J., & Vorbeck, M. (Eds.). (1982). *Social science research and public policy-making: A reappraisal.* Windsor, Berkshire: NFER-Nelson.

Kiesler, S.B., & Turner, C.F. (Eds.). (1977). *Fundamental research and the process of education.* Washington, DC: National Academy of Sciences.

Levy, P., & Healey, P. (1984). Education and human development. *ESRC Newsletter, 53*, 15-16. (Economic and Social Research Council, London).

Lindblom, C.E., & Cohen, D.K. (1979). *Usable knowledge.* New Haven: Yale University Press.

Marklund, I. (1981). *The impact of policy-oriented educational R & D.* Strasbourg, France: Council of Europe.

McDonald, G. (1980, November). *Proportion and perspective.* Address to the Second New Zealand Association for Research in Education Conference, Massey University.

McDonald, J. (1976, April). *Designing research for policy making.* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco (ERIC ED 128 370).

Mitchell, D.E. (1980). Social science impact on legislative decision making: Process and substance. *Educational Researcher, 9*(11), 9-19.

Mitchell, D.E. (1981). *Shaping legislative decisions: Educational policy and the social sciences*. Lexington, MA: Heath.

Nisbet, J., Megarry, J., & Nisbet, S. (Eds.). (1985). *World yearbook of education 1985: Research, policy, and practice*. London: Kogan Page.

Nisbet, J.D., & Broadfoot, P. (1980). *The impact of research on policy and practice in education*. Aberdeen: Aberdeen University Press.

Nuthall, G. (1975). Educational research in New Zealand. *Education, 24*(4), 3-7 (New Zealand Department of Education).

Paulston, C.B. (1982). *Swedish research and debate about bilingualism*. Stockholm: National Swedish Board of Education.

Powell, J.L. (1980, November). The impact of research. *Research in Education, 25*, 1-2 (Scottish Council for Research in Education).

Power, C. (1981). The contribution of research to educational policy and practice. *Australian Educational Researcher, 8*(1), 5-19.

Prewitt, K. (1981). Usefulness of the social sciences. *Science, 211*(4483), 659.

Radford, W.C., & McDowell, R. (1973). *Research into education in Australia 1972*. Canberra: Australian Government Printing Service.

Rothschild, Lord. (1982). *An enquiry into the Social Science Research Council*. London: Her Majesty's Stationery Office.

Rule, B.G. (1982, September). Objectives of social science research: Its perceived relevance and appropriate expectations. *Social Sciences of Canada, 10*(2), 2-4 (Social Science Federation of Canada).

Sheffield, E. (1982). *Research on postsecondary education in Canada*. Ottawa: Social Sciences and Humanities Research Council of Canada.

Shellard, J.S. (Ed.). (1979). *Educational research for policy making in Australia*. Hawthorn, Victoria: Australian Council for Educational Research.

Smith, S.L. (1984, March 7). *The challenge of the microprocessor to Canadian universities*. Address to the Association of Universities and Colleges of Canada, Montreal.

Social Science Federation of Canada. (1983). Task force report on the advancement of social science research in Canada in the 1980s: Consideration for the development of SSHRCC's new multiple-year plan. *Social Sciences in Canada, 11*(1), 2-6.

Steinle, J. (1981). Research and policy in education. *Proceedings of the Annual Conference of the Australian Association for Research in Education* (pp. 606-609). Adelaide: South Australian Government Printer.

Suppes, P. (Ed.). (1978). *Impact of research on education: Some case studies*. Washington, DC: National Academy of Education.

Supply and Services, Canada. (1985, January). Solving the drinking-driving dilemma: Government looks to research for help. *Research and Development Bulletin*, 142, ii-vi.

Taylor, W. (1973). *Research perspectives in education*. London: Routledge & Kegan Paul.

Taylor, W. (1981). Educational research and development in the United Kingdom: Framework, impact and future prospects. *International Review of Education, 27*(2), 179-195.

Trent, J.E. (1982). Promoting the social sciences: Lessons of four years. *Social Sciences in Canada, 10*(3), 2-4.

Wagenaar, H.C., Kallen, D.B.P., & Kosse, G.B. (1982). Social sciences and public policy-making in the 1980s: Major findings of the SVO workshop—Educational research and public policy-making. In D.B.P. Kallen, G.B. Kosse, H.C. Wagenaar, J.J.J. Kloprogge, & M. Vorbeck (Eds.), *Social science research and public policy-making: A reappraisal* (pp. 7-20). Windsor, Berkshire: NFER-Nelson.

Warren Piper, D. (1980). The application of theory in higher education. In R. Oxtoby (Ed.), *Higher education at the crossroads*. Proceedings of the 1980 Conference of the Society for Research in Higher Education (England).

Weiss, C.H. (Ed.). (1977). *Using social science research in public policy making*. Lexington, MA: Heath.

Weiss, C.H. (1982). Policy research in the context of diffuse decision making. *Journal of Higher Education, 53*(6), 619-639.

Wirt, F.M., & Mitchell, D.E. (1982). Social science and educational reform: The political uses of social research. *Educational Administration Quarterly, 18*(4), 1-16.

Wotjas, O. (1984, April 27). Social scientists 'want to play their part.' *Times Higher Education Supplement*, p. 4.

The Alberta Journal of Educational Research *Vol. XXXII, No. 3, September 1986, 265-267*

Social Sciences and Humanities Research Council of Canada

Strategic Grants Division
255 Albert Street
P.O. Box 1610
Ottawa, ON K1P 6G4

Education and Work in a Changing Society

The SSHRC offers grants for research on the relationship between education and work in a changing society under the above thematic program. For the purposes of the program, education encompasses more than simply schooling and refers rather to the lifelong process of learning, structured and informal.

At this time the Council is interested in receiving applications under the program within the broad areas outlined below. These are not mutually exclusive.

Skills—Generic and/or Specialized

The development of skills necessary for work, whether generic or specialized, as well as work habits and attitudes relating to the use of such skills, is an important aspect of the relationship between education and work. Examples of areas of interest that could be explored are:

- skills considered appropriate for work and the place where these are acquired—the home, educational institutions, the workplace, or elsewhere;
- the identification and development of skills (both direct and indirect) required to deal with new technologies;
- the place in the education system for the teaching of specific occupation skills;
- the appropriate role of various agencies in the development of skills—government, unions, educational institutions, employers, and others.

Stakeholders—Roles, Goles, and Returns

Many groups and organizations have an interest or stake in the relation between education and work and how this is articulated. Examples of these areas of interest that could be explored are:

- perceptions of the role and goals of the education system with regard to the world of work by parents, students, employers, governments, political parties, and ethnic and other groups or organizations;
- the role of education and other institutions in career change;

- the allocation of responsibility for education between and among governments, agencies, religious denominations, employers, the family, professional bodies, ethnic groups;
- standard setting by the education system and the labor market;
- concepts of returns on educational investment.

Responsiveness—Institutional and Systematic

The responsiveness of the education system and educational institutions to change is a major element in the relationship between education and work in a changing society. Examples of topics that could be explored are:

- mechanisms for response and adaptation;
- employment equity and education;
- the degree of linkage between the education system and labor market requirements;
- access to education and the role of education in occupational selection.

Education and Work in a Competitive International Environment

The ability to compete in an increasingly competitive international environment has been linked to an appropriate education system. Examples of topics that could be explored are:

- relevance of foreign educational models to Canada;
- relationships between Candian education and Canada's ability to compete internationally.

Programs of Support

1. Seed Money Grants for the Development of Proposals

   (Value: $5,000 maximum)

   The purpose of this program is to provide development funding to enable individuals or groups to expore initial formulation of project objectives, appropriate methodologies, time schedules, needed resources, and likely budget demands. It is expected tht these projects will culminate in a "mature" proposal for a Strategic Research Grant.
2. Strategic Research Grants

   (Value: $100,000 for one year/or $250,000 over three years)

   Grants are available for research projects, including multidisciplinary projects, in the Education and Work in a Changing Society field. Unless otherwise indicated, the regulations and conditions of Council's regular Research Grants Program also apply to the Strategic Research Grants Program. Applicants should therefore refer to the Research Grant's *Guide for Applicants* for detailed information on regulations and application requirements.
3. Research Workshops

   (Value not to exceed $15,000)

   The purpose of this program is to enable researchers to develop research skills and introduce themselves to new research ideas in the field of Education and

Work in a Changing Society. Awards are offered to assist in the organization and administration of workshops in which the participants can develop and exchange research skills and discuss new research needs.

Deadline for Applications

For 1987 awards, the deadline for receipt of applications is 1 April with results announced in September 1987.

Additional Information

Additional information is available from SSHRC Strategic Grants Division (address above).

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in triplicate and are not to exceed 25 pages in length.
2. An abstract of approximately 100 words in length, typed on a separate page, is to be included.
3. The author's name and affiliation should appear only on a separate title page to ensure anonymity in the reviewing process; the title of the paper should appear on the first page of the manuscript as a means of identification.
4. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publication Manual.
5. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
6. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevant point the phrase (Insert Table 2 here).
7. Sources should appear in parentheses following the reference, citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References.*
8. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes.* They should be double spaced and placed at the end of the manuscript immediately preceding the *References.* The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 7 above.
9. Spelling shall conform to *Webster's New Collegiate Dictionary*, except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
10. In matters of style, the third edition of the APA Publication Manual is considered definitive.
11. Submission implies that the manuscript is not being considered for publication elsewhere.

The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER is indexed in the *Alberta Education Index, Canadian Education Index, Current Contents/Social and Behavioral Sciences, ERIC/Current Index to Journals in Education,* and *Social Science Citation Index;* appropriate articles are abstracted in *Cognitive Development Abstracts, Educational Administration Abstracts, Educational Technology Abstracts, Family Abstracts, Language Behavior Abstracts, Multicultural Education Abstracts, Psychological Abstracts, Research into Higher Education Abstracts, School Organization and Management Abstracts,* and *Sociology of Education Abstracts,* and reviewed by the *Journal for Research in Mathematics Education.*